

Die Wirksamkeit bilingualen Sachfachunterrichts:

Selektionseffekte, Leistungsentwicklung und die Rolle der Sprachen
im deutsch-englischen Geschichtsunterricht

Dissertation
zur Erlangung des Grades eines Doktors
der Philosophie (Dr. phil.)

der Pädagogischen Hochschule Ludwigsburg

vorgelegt von Sara Dallinger, geb. Diezis, aus Ludwigsburg

Ludwigsburg
2015

Erstgutachter: Prof. Dr. Jan Hollm
Zweitgutachterin: Prof. Dr. Kathrin Jonkmann
Drittgutachter: Prof. Dr. Jörg-U. Keßler
Viertgutachter: Prof. Dr. Götz Schwab

Datum des Abschlusses der mündlichen Prüfung: 10. Juni 2015

DANKSAGUNG

Die vorliegende Arbeit entstand im Rahmen meines Promotionsstipendiums und innerhalb des Kooperativen Promotionskollegs „Effektive Lehr-Lernarrangements: Empirische Evaluation und Intervention in der pädagogischen Praxis“ zwischen der Pädagogischen Hochschule Ludwigsburg und der Universität Tübingen. Hierbei möchte ich mich bei allen Beteiligten, insbesondere bei Prof. Dr. Ulrich Trautwein als Sprecher des Kollegs, für die Unterstützung und für diese einmalige Chance zum Austausch und zur Vernetzung im vielfältigen Bereich der empirischen Bildungsforschung bedanken.

Ganz besonderer Dank gilt meinen beiden Betreuern, die mir bei der Entstehung dieser Arbeit immer sehr unterstützend zur Seite standen: Prof. Dr. Kathrin Jonkmann, bei der ich mich darüber hinaus für ihre unglaubliche Expertise bedanken möchte, die mich so viel weitergebracht hat; und bei Prof. Dr. Jan Hollm, dessen unermüdliche fachliche und persönliche Unterstützung und dessen Glaube an mich all dies erst ermöglicht hat.

Des Weiteren danke ich herzlich Prof. Dr. Jörg-U. Keßler, Prof. Dr. Bärbel Völkel, Christiane Bertram und Dr. Christiane Fiege für die vielfältige Unterstützung meiner Arbeit. Nicole Piesche danke ich für unsere netten Unterhaltungen unter Gleichgesinnten. Weiterhin vielen Dank an Andrea Hettler und meine anderen studentischen Hilfskräfte, deren tatkräftige und zuverlässige Hilfe die (fast) reibungslose Umsetzung meiner Forschungsarbeit gewährleistete. Schließlich gilt ein großer Dank Kirsten Riegler, auf deren ausgezeichnete und zuverlässige Unterstützung ich mich stets verlassen durfte.

Diese Arbeit ist meiner Familie gewidmet – meinen Eltern und meinen Brüdern – ohne die all dies bedeutungslos wäre. Und Boris.

INHALT

| | | |
|-----|--|-----|
| 1 | EINLEITUNG UND THEORETISCHER RAHMEN DER ARBEIT | 1 |
| 1.1 | Das Angebots-Nutzungs-Modell als theoretische Grundlage zur Analyse der Wirksamkeit bilingualen Unterrichts..... | 7 |
| 1.2 | Die Schülerklientel des bilingualen Unterrichts | 13 |
| 1.3 | Die bilingual unterrichtenden Lehrkräfte | 18 |
| 1.4 | Das bilinguale Unterrichtsangebot..... | 22 |
| 1.5 | Schülerleistungen im bilingualen Unterricht | 28 |
| 1.6 | Bisherige Forschung zur Wirksamkeit bilingualen Unterrichts..... | 38 |
| 1.7 | Fragestellungen der vorliegenden Arbeit..... | 51 |
| 1.8 | Literatur | 60 |
| 2 | COMPETENCES AND MOTIVATION IN BILINGUAL INSTRUCTION IN HISTORY (COMBIH) – EINE LÄNGSSCHNITTSTUDIE ZU DEUTSCH-ENGLISCHEM GESCHICHTSUNTERRICHT AN GYMNASIEN..... | 73 |
| 3 | LERNAUSGANGSLAGEN VON SCHÜLERINNEN UND SCHÜLERN ZU BEGINN DES BILINGUALEN GESCHICHTSUNTERRICHTS | 93 |
| 4 | SELECTION EFFECTS IN CONTENT AND LANGUAGE INTEGRATED LEARNING CLASSROOMS IN GERMAN SECONDARY SCHOOLS | 105 |
| 5 | THE EFFECT OF CONTENT AND LANGUAGE INTEGRATED LEARNING ON STUDENTS’ ENGLISH AND HISTORY COMPETENCES – KILLING TWO BIRDS WITH ONE STONE? | 123 |
| 6 | MERKMALE ERFOLGREICHEN BILINGUALEN SACHFACHUNTERRICHTS – DIE ROLLE DER SPRACHEN IM DEUTSCH-ENGLISCHEN GESCHICHTSUNTERRICHT | 145 |
| 7 | ABSCHLUSSDISKUSSION | 167 |
| 7.1 | Zusammenfassung und Bedeutung der zentralen Befunde | 168 |
| 7.2 | Forschungsperspektiven | 177 |
| 7.3 | Fazit | 184 |
| 7.4 | Literatur | 185 |

ZUSAMMENFASSUNG

Bilingualer Sachfachunterricht wird in Deutschland bereits seit 1969 in Form fest etablierter Züge an Gymnasien angeboten. Insbesondere seit dem Jahrtausendwechsel erlebte diese Unterrichtsform einen regelrechten Boom und wurde in den meisten Bundesländern systematisch und flächendeckend ausgebaut (KMK, 2013: 3-5). Parallel zu diesen Entwicklungen entstand, überwiegend im Bereich der Fremdsprachendidaktik, eine rege Forschungstätigkeit, die zu zahlreichen Publikationen führte. Dennoch bedarf es nach wie vor an inhaltlich umfassenden und forschungsmethodisch adäquat durchgeführten Studien zur Wirksamkeit bilingualen Unterrichts (Rumlich, 2014a; Bonnet, 2012; Pérez-Cañado, 2012; Bruton, 2011; Costa & D'Angelo, 2011). Dies stellt das Ziel der vorliegenden Arbeit zu deutsch-englischem bilingualen Geschichtsunterricht dar. Hierfür fand in Anlehnung an Helmkes (2012) Angebots-Nutzungs-Modell eine interdisziplinäre Zusammenarbeit zwischen der Fremdsprachen- und Sachfachdidaktik sowie der empirischen Bildungsforschung statt. Mithilfe der Integration bisheriger Erkenntnisse, spezifischer Überlegungen und methodischer Vorgehensweisen der jeweiligen Wissenschaftsdomänen wurden die Effekte dieser Unterrichtsform auf die Entwicklung der Englisch- sowie Geschichtskompetenzen der bilingual unterrichteten Lerner unter Berücksichtigung verschiedenster Schüler- und Klassenmerkmale untersucht sowie die Wirksamkeit spezifischer methodisch-didaktischer Unterrichtsaspekte in Bezug auf die Schülerleistungen bestimmt. Dies erfolgte innerhalb der folgenden empirischen Teilstudien.

Teilstudie I.A untersuchte mögliche Unterschiede in den Lernausgangslagen zwischen 703 bilingual und 1103 einsprachig unterrichteten Achtklässlern zu Beginn des bilingualen Geschichtsunterrichts, die sich aufgrund schulischer Auswahlprozesse im Zusammenhang mit dem Besuch eines bilingualen Zugs ergeben haben könnten (Zydati, 2007: 70; Bredenbröcker, 2000: 50). Solche unterschiedlichen Voraussetzungen müssen bei längsschnittlichen Untersuchungen zur Wirksamkeit des bilingualen Unterrichts berücksichtigt werden, um dessen Effekte bestmöglich isolieren zu können. Hierfür wurden die Daten der bilingual unterrichteten Lerner mit den Daten von zwei Vergleichsgruppen verglichen: einsprachig unterrichtete Schülerinnen und Schüler aus den Parallelklassen der Bilingualklassen sowie aus Schulen ohne bilinguales Unterrichtsangebot. Die Analysen ergaben Vorteile zugunsten der bilingual unterrichteten in ihrem sozioökonomischen Status, den verbalen kognitiven Grundfähigkeiten sowie in der Motivation und im Vorwissen in Englisch und in Geschichte. Bezüglich figuraler kognitiver Grundfähigkeiten zeigten sich keine statistisch bedeutsamen Unterschiede zwischen den Schülerinnen und Schülern. Dasselbe traf auf kleinere Unterschiede zwischen den beiden einsprachig unterrichteten Vergleichsgruppen zu.

Teilstudie I.B befasste sich anknüpfend an die Befunde aus Teilstudie I.A mit möglichen Selektionseffekten bilingualer Züge und beantwortete die Frage, ob sich die erwarteten Leistungsvorteile bilingual unterrichteter Schülerinnen und Schüler in Englisch durch deren möglicherweise günstigeren Lernvoraussetzungen erklären lassen (Rumlich, 2013: 184; Bruton,

2011). Hierfür wurden die in Teilstudie I.A analysierten Aspekte sowie weitere Unterschiede zwischen den Schülergruppen bezüglich der Bildungsnähe des Elternhauses berücksichtigt. Die Analyse der Daten der 703 bilingual und 659 einsprachig unterrichteten Lerner aus Schulen mit bilinguaem Zug ergab, dass sich die Vorteile der bilingual Unterrichteten in den Englischleistungen deutlich verringerten, was auf Selektionseffekte zu deren Gunsten verweist. Es bestanden aber dennoch weiterhin Vorteile in Höhe von mindestens eineinviertel Schuljahren fort.

Teilstudie II untersuchte auf Basis von verschiedenen Zweitspracherwerbtheorien und kognitionspsychologischen Erkenntnissen (Adesope, Lavin, Thompson & Ungerleider, 2010; Gass & Mackey, 2007; Krashen & Terrell, 2000; Krashen, 1985; Cummins, 1979) die Kompetenzentwicklung von 30 bilingual unterrichteten Klassen (703 Lerner) in Englisch und Geschichte im Verlauf der 8. Klassenstufe und verglich diese mit der Kompetenzentwicklung von 45 einsprachig unterrichteten Schülergruppen (1103 Lerner). Hierbei wurden in Anknüpfung an die Befunde aus den Teilstudien I.A und I.B mögliche Unterschiede zwischen bilingual und einsprachig unterrichteten Klassen berücksichtigt, und zwar in den durchschnittlichen Vorkenntnisständen, klassenbezogenen Lernvoraussetzungen sowie in der Unterrichtsqualität in Englisch und Geschichte und in Lehrermerkmalen. Die Leistungsentwicklung der bilingual unterrichteten Gruppen verlief im Hörverstehen deutlich positiver als bei den einsprachig unterrichteten Klassen. Bilingual Unterrichtete verbesserten diese Englischkompetenz also merklich stärker. Die Entwicklung der allgemeinen englischen Sprachfähigkeit sowie die Aneignung von Sachwissen in Geschichte hingen im Verlauf des Schuljahres vergleichbar zwischen den Schülergruppen.

Teilstudie III befasste sich in Anlehnung an Prinzipien und Theorien des Fremdspracherwerbs (Gass & Mackey, 2007; Krashen, 1985; Cummins, 1979; Butzkamm, 1973) mit der Wirksamkeit des Einsatzes von Mutter- und Fremdsprache hinsichtlich der Schülerkompetenzen im bilingualen Unterricht. Somit wurde untersucht, wie spezifische methodisch-didaktische Merkmale dieser Unterrichtsform die Leistungsentwicklung der bilingual unterrichteten Klassen bedingten unter gleichzeitiger Kontrolle von sozioökonomischem Status, Motivation und Vorkenntnisstand im Klassendurchschnitt. Die Auswertung der längsschnittlich erhobenen Daten von 30 bilingual unterrichteten achten Klassen (703 Lerner) und deren Lehrkräften ergab, dass ein verstärkter Englischinput durch die Lehrkraft mit günstigen Entwicklungen der Englisch- sowie Geschichtskompetenzen der Schülergruppen einherging. Dasselbe traf auf den planvollen Miteinbezug der Muttersprache, insbesondere bei der Einführung neuer Fachtermini, zu. Ein großer englischer Redeanteil seitens der Schülerinnen und Schüler hingegen ging mit negativeren Leistungsentwicklungen einher.

In der Abschlussdiskussion wurden die Ergebnisse der vorliegenden Arbeit zusammengefasst und deren forschungsmethodisch bedingte Bedeutung aufgezeigt. Anschließend wurden Implikationen für die Schul- und Unterrichtspraxis abgeleitet sowie Forschungsperspektiven für zukünftige Untersuchungen zur Wirksamkeit bilingualen Unterrichts eröffnet.

Einleitung und theoretischer
Rahmen der Arbeit

1 Einleitung und theoretischer Rahmen der Arbeit

*“Language is not an end in itself, just as little as railway tracks;
it is a way of connection between souls, a means of communication.”*

(Jespersen, 1904: 3)

In den letzten 15 Jahren erfuhr bilingualer Sachfachunterricht, bei dem Schulfächer wie beispielsweise Geschichte oder Geographie überwiegend in einer Fremdsprache sowie anteilig in der Muttersprache gelehrt werden, einen Aufschwung in Deutschland und wurde systematisch ausgebaut, so dass diese Unterrichtsform inzwischen in allen Bundesländern in den verschiedensten Schulformen, Fächern und Fremdsprachen angeboten wird (KMK, 2013: 11-14). Auch auf europäischer Ebene erlebte bilingualer Unterricht (bzw. *Content and Language Integrated Learning*; CLIL), mit dem sowohl Fremdsprachen- als auch Sachfachkompetenzen der Schülerinnen und Schüler gefördert werden sollen (KMK, 2013: 5-7; Eurydice, 2006: 22), einen Boom und hat sich mittlerweile als fester Bestandteil des Unterrichtsangebots in fast allen Ländern etabliert (Pérez-Cañado, 2012: 319). Diese Entwicklungen gingen mitunter auf verschiedene EU-Beschlüsse Mitte der 1990er Jahre und Anfang des neuen Jahrtausends zurück, die – angesichts der zunehmenden Globalisierung und der Öffnung europäischer Grenzen – die Mehrsprachigkeit von Schülerinnen und Schülern forderten (Eurydice, 2006: 8; KMK, 2006: 8). Folglich entstanden insbesondere zu Beginn des neuen Jahrtausends zahlreiche Forschungsarbeiten (z. B. Bonnet, 2004; Hermes & Klippel, 2003; Wildhage & Otten, 2003; Breidbach, Bach & Wolff, 2002; Daniel, 2001), die diese Entwicklungen wissenschaftlich begleiteten. Diese und weitere Studien konnten bisher zeigen, dass sich bilingualer Sachfachunterricht positiv auf die Fremdsprachenkompetenzen (Köller, Leucht & Pant, 2012; Nold, Hartig, Hinz & Rossa, 2008; Zydati, 2007) sowie auf die Motivation (Rumlich, 2014a; Abendroth-Timmer, 2007) der Schülerinnen und Schüler in der Fremdsprache auszuwirken schien. Die wenigen vorliegenden Arbeiten zu den Leistungen der Lerner im bilingual unterrichteten Sachfach verwiesen auf vergleichbare Ergebnisse der zwei- und einsprachig unterrichteten Schülerinnen und Schüler (Badertscher & Bieri, 2009; Jäppinen, 2005; Bonnet, 2004). Darüber hinaus scheint der bilinguale Unterricht Unterstützung von der Mehrheit der Lehrenden (Hüttner & Dalton-Puffer, 2013; Kollenrott, 2008; Viebrock, 2007; Müller-Schneck, 2006), der Elternschaft (Dallinger, 2013) sowie seitens der Schulverwaltungen (Hollm, 2013) zu erfahren.

Trotz dieser Entwicklungen und der Vielzahl an Untersuchungen, die auf eine „Erfolgsgeschichte“ des bilingualen Unterrichts hinweisen, kann argumentiert werden, dass die Wirksamkeit dieser Unterrichtsform hinsichtlich der Schülerleistungen in der Fremdsprache und im Sachfach noch nicht umfassend, und in manchen Fällen unter Inkaufnahme forschungsmethodischer Schwächen, analysiert wurde und deshalb weiterhin eine offene Frage bleibt. Dies spiegelt sich in

diverser Kritik wieder, die insbesondere in den letzten Jahren vermehrt geäußert wurde (Rumlich, 2014a; Bonnet, 2012; Pérez-Cañado, 2012; Bruton, 2011). Bruton (2011) argumentierte beispielsweise, dass bisherige Studien nur selten die Ausgangslagen der Lerner vor Beginn des bilingualen Unterrichts analysierten. Angesichts der von ihm festgestellten hohen Selektivität bilingualer Unterrichtsprogramme vermutete er, dass bilingual unterrichtete Lerner im Vergleich zu einsprachig unterrichteten Schülerinnen und Schülern über besonders günstige Lernvoraussetzungen verfügen könnten, z. B. hinsichtlich ihrer kognitiven Grundfähigkeiten oder ihrer (fremdsprachenbezogenen) Lernmotivation. Dies wirke sich Bruton zufolge positiv auf deren Schulleistungen aus und münde in einem grundsätzlich höheren Fremdsprachenniveau der bilingual unterrichteten Lerner bzw. bilingual unterrichteten Klassen im Vergleich zu einsprachig unterrichteten Schülerinnen und Schülern. Folglich scheint gegenwärtig unklar, ob Vorteile in den Kompetenzen, die in verschiedenen Studien zugunsten bilingual Unterrichteter gefunden wurden, durch den bilingualen Unterricht bedingt wurden oder – als Resultat möglicher Selektionsprozesse – auf ihre günstigeren Lernvoraussetzungen und größeren Vorkenntnisse zurückzuführen sind (Bonnet, 2012: 67). Weiterhin argumentierte Rumlich (2014a), dass diese fehlenden Analysen der Ausgangslagen von bilingual und einsprachig unterrichteten Schülerinnen und Schülern auch denjenigen Faktor darstellen, aufgrund dessen die meisten querschnittlichen Studien kaum Aussagen zur Wirksamkeit bilingualen Unterrichts treffen können. Denn da diese die Vorkenntnisse der Lerner nicht berücksichtigen konnten, scheinen in den meisten Querschnittsstudien die Effekte bilingualen Unterrichts auf die Schülerleistung systematisch überschätzt worden zu sein (ibid.: 83).

Darüber hinaus gibt es weitere Aspekte, die in Wirksamkeitsanalysen zu bilinguaem Unterricht berücksichtigt werden müssen, da sie die Schülerleistung mit beeinflussen, jedoch nicht in direktem Zusammenhang mit dem bilingualen Unterrichtsgeschehen stehen. Solche Aspekte müssen statistisch kontrolliert werden, um die tatsächlichen Effekte dieser Unterrichtsform bestmöglich isolieren zu können. Dies trifft beispielsweise auf die bilingual bzw. einsprachig unterrichtende Lehrperson sowie die damit in Zusammenhang stehende Unterrichtsqualität zu, die sich zwischen bilingual und einsprachig unterrichteten Klassen systematisch unterscheiden könnten (Dalton-Puffer, Nikula & Smit, 2010: 282). Bisher sind jedoch keine Studien bekannt, die die Wirksamkeit bilingualen Unterrichts unter Berücksichtigung all dieser Aspekte untersucht haben. Es bedarf also Längsschnittstudien (Costa & D'Angelo, 2011: 3), die die Leistungsentwicklung bilingual und einsprachig unterrichteter Klassen analysieren und hierbei Unterschiede in den durchschnittlichen Vorkenntnisständen, Lernvoraussetzungen und in weiteren klassenbezogenen Merkmalen, wie Lehrperson und Unterrichtsqualität, statistisch kontrollieren. Hierbei ist zudem die Beschaffenheit der Vergleichsgruppe zu beachten. Denn bisher wurden für vergleichende Analysen zwischen bilingual und einsprachig unterrichteten Lernern überwiegend solche einsprachig unterrichteten Schülerinnen und Schüler herangezogen, die dieselben Schulen wie die bilingual Unterrichteten besuchten. Brutons (2011) Logik folgend könnten diese jedoch im Vergleich zur

bilingualen Lernergruppe eine negativ selektierte Schülerschaft und somit eine verzerrte Vergleichsgruppenstichprobe darstellen (Rumlich, 2012: 173). Deshalb sollte in Wirksamkeitsstudien zu bilinguaem Unterricht eine weitere Vergleichsgruppe hinzugezogen werden, die sich aus Lernern von Schulen ohne bilinguales Unterrichtsangebot zusammensetzt.

Abschließend besteht ein weiteres Forschungsdesiderat in Bezug auf die Frage, welche spezifischen methodisch-didaktischen Aspekte im bilingualen Unterricht zur Ausbildung dieser Schülerleistungen führen könnten. Hierbei handelt es sich um die Wirksamkeit bilingualen Unterrichts, die nicht im Vergleich zu einsprachigem Sachfachunterricht bestimmt wird, sondern die sich auf Wirkungsweisen innerhalb des bilingualen Unterrichtssettings bezieht. Es bedarf also Studien, die mögliche Effekte spezifischer Unterrichtsmaßnahmen auf die Leistungen der bilingual unterrichteten Klassen in den Mittelpunkt rücken. In diesem Zusammenhang wären beispielsweise Art und Umfang des Einsatzes von Mutter- und Fremdsprache bedeutsam; also derjenige Aspekt, der bilingualen von regulär einsprachigem Sachfachunterricht unterscheidet. Trotz der Aktualität der Debatte zur Rolle der Sprachen im bilingualen Unterricht (vgl. Diehr & Schmelter, 2012) blieb dieser Aspekt bislang weithin unerforscht (Diehr, 2012: 29).

Die eben aufgezeigten Forschungsdesiderate in Bezug auf die Wirksamkeit bilingualen Unterrichts mögen mitunter darin begründet liegen, dass für solche umfassenden Analysen bisher kaum interdisziplinär zusammen gearbeitet wurde. Im Idealfall bedürfte es der integrativen Expertise von Fachdidaktikern sowie Vertretern der empirischen Bildungsforschung, um sowohl inhaltlich den entsprechenden Fragestellungen als auch formal den forschungsmethodischen und statistisch-analytischen Anforderungen gerecht zu werden. Solche interdisziplinären Forschungsvorhaben wurden im Kontext des bilingualen Unterrichts jedoch bisher kaum umgesetzt. Weiterhin bedarf es einer solchen Zusammenarbeit auch zwischen den Fremdsprachen- und Sachfachdidaktiken. Bisherige Untersuchungen nahmen entweder die fremdsprachliche Perspektive auf den bilingualen Unterricht (z. B. Rumlich, 2014b; Fehling, 2008) oder, deutlich seltener, Aspekte im Zusammenhang mit dem bilingual unterrichteten Sachfach in den Blick (Bonnet, 2012: 67; Bernhardt, 2011: 215; Coyle, Hood & Marsh, 2010: 138). Dementgegen wurden jedoch kaum Studien zur Wirksamkeit bilingualen Unterrichts durchgeführt, die sowohl Sprach- als auch Sachfachkompetenzen der Lerner berücksichtigten.

Auf Grundlage dieser Überlegungen beschäftigt sich die vorliegende Arbeit zu deutsch-englischem Geschichtsunterricht in ihren vier empirischen Teilstudien mit folgenden Fragen: Bestehen bereits zu Beginn des bilingualen Geschichtsunterrichts Unterschiede in den Lernvoraussetzungen und Vorkenntnissen in Englisch und Geschichte zugunsten der bilingual unterrichteten Schülerinnen und Schüler und bestehen solche Unterschiede auch zwischen einsprachig unterrichteten Lernern von Schulen mit und ohne bilinguaem Unterrichtsangebot (Teilstudie I.A)? Können mögliche Vorteile der bilingual unterrichteten Schülerinnen und Schüler in Englisch durch deren günstigere Lernvoraussetzungen erklärt werden (Teilstudie I.B)?

Unterscheidet sich die Kompetenzentwicklung von bilingual und einsprachig unterrichteten Klassen in Englisch und Geschichte unter statistischer Kontrolle verschiedener Klassenmerkmale (Teilstudie II)? Welche Wirksamkeit zeigen der Einsatz von Mutter- und Fremdsprache im bilingualen Unterricht hinsichtlich der Leistungsentwicklung bilingual unterrichteter Klassen (Teilstudie III)?

Zur angemessenen Beantwortung dieser Fragen fand im Rahmen der vorliegenden Studie eine interdisziplinäre Zusammenarbeit zwischen Sprach- und Sachfachdidaktiken sowie empirischer Bildungsforschung statt, wobei die Expertise aller drei Wissenschaftsbereiche integriert wurde: 1. Erkenntnisse aus dem Bereich der empirischen Bildungsforschung zur Entwicklung des Studiendesigns, allgemein bildungspsychologischer Fragestellungen und zur Durchführung verschiedener psychometrischer und statistisch-analytischer Verfahren sowie fachdidaktische Expertise in 2. Englisch und 3. Geschichte zur Formulierung fachspezifischer Fragestellungen, der Entwicklung der Leistungstests und zur Erfassung methodisch-didaktischer Unterrichtsaspekte. Diese Interdisziplinarität verfolgte den Ansatz einer möglichst ganzheitlichen und validen Erfassung des vorgefundenen Schulsettings und trägt zur Annäherung an das Forschungsziel der vorliegenden Arbeit bei, nämlich das Gewinnen von belastbaren Erkenntnissen zur Wirksamkeit bilingualen Geschichtsunterrichts in Bezug auf die Leistungsentwicklung der Schülerinnen und Schüler sowie das Leisten eines Beitrags zur Entwicklung einer empirisch fundierten Methodik dieser Unterrichtsform.

Der Aufbau der vorliegenden publikationsorientierten und deshalb kumulativ organisierten Dissertation gestaltet sich wie folgt: Die Einleitungskapitel (Kapitel 1 und 2) stellen den theoretischen und empirischen Rahmen dar, innerhalb dessen sich die Teilstudien (Kapitel 3 bis 6) verorten lassen. Abschnitt 1.1 stellt zunächst das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts nach Helmke (2012) vor. Das Modell dient als theoretische Basis für die möglichst umfassende und interdisziplinäre Erhebung des bilingualen Unterrichts und seiner beteiligten Personen und Prozessmerkmale. Es wird konkretisiert, auf welche Weise dieses Modell in der vorliegenden Arbeit für die Analyse der Wirksamkeit dieser Unterrichtsform zum Einsatz kam. Daran anknüpfend vertiefen Abschnitte 1.2 bis 1.5 diejenigen Aspekte des Angebots-Nutzungs-Modells, die für den Kontext der vorliegenden Arbeit als relevant identifiziert wurden. Der Fokus wird hierbei jeweils auf der inhaltlichen Ableitung und Konkretisierung dieser Kategorien liegen. So behandelt Abschnitt 1.2 die Aspekte „Familie“ und „Lernpotential“ des Modells, die in Wechselwirkung mit dem „Ertrag“, also den Kompetenzen der Lerner, stehen. In diesem Zusammenhang wird die Organisation bilingualer Züge vorgestellt und die daraus resultierende spezifische Schülerklientel des bilingualen Unterrichts, mit ihren möglicherweise günstigeren Lernvoraussetzungen, thematisiert. Darüber hinaus wird auch der (Unterrichts-)„Kontext“ dem Abschnitt 1.2 zugeordnet. Er bezieht sich in der vorliegenden Arbeit vornehmlich auf die Klassenzusammensetzung der bilingual unterrichteten Lernergruppen und überträgt die Frage

möglicher günstigerer Lernvoraussetzungen von der Schüler- auf die Klassenebene. Abschnitt 1.3 befasst sich mit der „Lehrperson“, also mit bilingual unterrichtenden Lehrkräften, und legt dar, inwieweit sich diese von ihren regulär einsprachig unterrichtenden Kolleginnen und Kollegen unterscheiden könnten. In Abschnitt 1.4 wird der Aspekt des (Unterrichts-),„Angebots“ behandelt. Dies geschieht in der vorliegenden Arbeit mithilfe zweier Perspektiven. Aus fachdidaktischer Perspektive werden spezielle, auf die Unterrichtssprache bezogene, methodisch-didaktische Maßnahmen thematisiert, die als Spezifikum des bilingualen Unterrichts zum Einsatz kommen. Die zweite Perspektive auf das Unterrichtsangebot stellt eine allgemein schulpädagogische Sicht dar. Sie nimmt das Konzept der Unterrichtsqualität in den Blick und adressiert diesbezüglich mögliche Unterschiede zwischen bilingual und einsprachig unterrichteten Klassen. Schließlich wird thematisiert, inwieweit sich die individuelle Wahrnehmung des Unterrichtsangebots durch Leistungsemotionen zwischen den bilingual und einsprachig unterrichteten Schülerinnen und Schülern unterscheiden könnte. Abschnitt 1.5 befasst sich mit dem „Ertrag“-Aspekt des Modells. Zunächst wird eine Verbindung zwischen Unterrichtsangebot und diesem Ertrag hergestellt, um die theoretische Fundierung des bilingualen Unterrichts aufzuzeigen. Hierzu werden Prinzipien und Modelle der Englisch- und Geschichtsdidaktik sowie Theorien aus dem Bereich des Zweitsprachenerwerbs und kognitionspsychologische Erkenntnisse vorgestellt und auf den Kontext des bilingualen Sachfachunterrichts angewandt. Danach wird das Konzept der „Schülerleistung“ bzw. „Schülerkompetenz“ näher betrachtet und domänenspezifisch für die vorliegende Arbeit definiert. Abschnitt 1.6 stellt im Anschluss daran die bisherige Forschung zur Wirksamkeit bilingualen Sachfachunterrichts vor und konzentriert sich hierbei auf Studien, die sich mit der Leistungsentwicklung der Lerner beschäftigten sowie mit Aspekten, die diese möglicherweise bedingen. In Abschnitt 1.7 werden schließlich die Fragestellungen der vorliegenden Arbeit aus den Darstellungen der vorhergehenden Abschnitte abgeleitet und unter Bezugnahme auf die empirischen Teilstudien konkretisiert, welche dann in den Kapiteln 3 bis 6 dargestellt werden. Zuvor erfolgt in Kapitel 2 die detaillierte konzeptionelle und forschungsmethodische Darstellung der vorliegenden Arbeit mit Fokus auf deren interdisziplinär-triangulativen Gehalt. Kapitel 7 stellt eine die Teilstudien umfassende Abschlussdiskussion dar, in der die zentralen Befunde zusammengefasst, die Implikationen der Studie verdeutlicht sowie ihre Grenzen aufgezeigt und Impulse für zukünftige Forschung gesetzt werden.

1.1 Das Angebots-Nutzungs-Modell als theoretische Grundlage zur Analyse der Wirksamkeit bilingualen Unterrichts

1.1.1 Das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts

Die vorliegende Arbeit beschäftigt sich mit der Wirksamkeit des bilingualen Unterrichts. Hierbei steht vor allem die Analyse von Effekten dieser Unterrichtsform auf die Leistungsentwicklung der Schülerinnen und Schüler im Vordergrund. Um dies leisten zu können, bedarf es eines theoretischen Rahmens, der einerseits Schülerkompetenzen als eine Wirkung des Unterrichtsgeschehens begreift und andererseits dieses Unterrichtsgeschehen – im vorliegenden Fall der bilinguale Unterricht – über verschiedene theoretische Aspekte abbildet und somit zugänglich macht. Dies leistet das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts nach Helmke (2012), weshalb es für die vorliegende Arbeit als theoretische Basis für die möglichst umfassende und interdisziplinäre Erhebung des bilingualen Unterrichts und seiner beteiligten Personen und Prozessmerkmale dient.

Das Modell beschreibt Determinanten von Schulleistung auf Mikroebene, konzentriert sich also hauptsächlich auf Kernelemente wie Unterricht, Lehrkraft und Lerner. Somit ist es Bestandteil eines umfassenderen Makromodells der Bedingungsfaktoren schulischer Leistung (Helmke, 2012: 28), das auch andere (außerschulische) Einflussfaktoren, wie beispielsweise Medien, Persönlichkeit der Eltern und Gleichaltrige, umfasst. Das Angebots-Nutzungs-Modell basiert sowohl auf theoretischen Überlegungen von Fend (1981) und Helmke und Weinert (1997) als auch auf empirischer Evidenz (Helmke, 2012: 29f.). Es geht davon aus, dass Lehrkräfte ihren Schülerinnen und Schülern ein (Unterrichts-)Angebot machen. Dies führt jedoch nicht per se zu Lernerfolg, sondern muss von den Lernern aktiv genutzt werden. Entsprechend dieses konstruktivistischen Verständnisses von Lernen stellen sich Wirkungen seitens der Schülerinnen und Schüler in Abhängigkeit davon ein, ob das Angebot überhaupt wahrgenommen und wie es interpretiert wird und zu welchen emotionalen, motivationalen und volitionalen Prozessen es bei den Lernern führt. Das Angebots-Nutzungs-Modell umfasst sechs Kategorien, die Einfluss auf Art und Umfang der „Wirkung“ nehmen, also auf den Ertrag seitens der Lerner, z. B. hinsichtlich ihrer Kompetenzen. Die sechs Kategorien lauten „Kontext“, „Familie“, „individuelles Lernpotential“, „Lehrperson“, „Unterricht“ und „Lernaktivitäten“, welche nun im Folgenden näher vorgestellt werden.

Kontext. Nach Helmke lässt sich der Unterrichtskontext in sechs Aspekte aufgliedern. 1. Die kulturellen Rahmenbedingungen beeinflussen das Unterrichtsgeschehen durch die sozialisatorische Funktion, die die Gesellschaft im Allgemeinen sowie die Schule im Konkreten übernimmt, indem beispielsweise technisches Wissen oder die Leistungsideologie eines Kulturkreises vermittelt wird. 2. Der regionale Kontext verweist auf kontextuelle Unterschiede innerhalb Deutschlands und kann sich sowohl auf schulinstitutionelle Aspekte (z. B. Vorhandensein von Gesamtschulen in verschiedenen Bundesländern) als auch auf ökonomische (z. B. Süd-Nord-Gefälle) oder

soziologische Faktoren (z. B. Kriminalitätsrate in Landkreisen) beziehen. 3. Ein weiterer kontextueller Aspekt stellt die Schulform bzw. der Bildungsgang dar. Dabei ist von Bedeutung, ob sich Lerner beim Unterricht beispielsweise in einer Grund-, Haupt-, Real- oder Förderschule oder auf einem Gymnasium befinden. 4. Die Klassenzusammensetzung beeinflusst nach Helmke das Unterrichtsgeschehen ebenfalls. Hierzu zählen z. B. die Klassengröße, das Vorkenntnisniveau der Klasse sowie die Schichtzusammensetzung und der Anteil der Schülerinnen und Schüler der Klasse mit Migrationshintergrund. 5. Der didaktische Kontext, d. h. fachspezifische Prinzipien, Strategien und Methoden des Lehrens und Lernens, spielen naturgemäß ebenfalls eine Rolle im Unterricht. 6. Abschließend tragen auch Schul- und Klassenklima zur Wirkung des Unterrichts bei. Dies trifft sowohl auf Verhaltensabläufe in der Schule im Allgemeinen zu (Umgang zwischen Lehrkräften, Lernern, Eltern und Schulleitung; Achtung und Einhaltung der Schuldisziplin) als auch auf Werte und Empfindungen innerhalb von Klassen (Kooperation, Vertrauen, Kameradschaft, Identifikation mit der Unterrichtsarbeit, Wettbewerbsdruck).

Familie. Die strukturellen Merkmale einer Familie stellen Aspekte dar, die in Wechselwirkung mit dem individuellen Lernpotential eines Lerners stehen. Solche familiären Merkmale sind nach Helmke beispielsweise soziale Schichtzugehörigkeit, zuhause gesprochene Sprache, Kulturzugehörigkeit, Bildungsnähe, Anregungsgehalt der Lernumwelt, Familienstruktur sowie Prozessmerkmale der Erziehung und Sozialisation (inklusive elterliches Engagement und Erwartungen). Sie wirken sich auf die Quantität und Qualität der Nutzung des Unterrichtsangebots durch den Lerner aus und tragen somit indirekt zu Art und Umfang des möglichen schulischen Ertrags bei.

Lernpotential. Das individuelle Lernpotential eines Kindes hängt von vielfältigen Aspekten ab. Nach Helmke sind dies insbesondere Vorkenntnisse, Sprachkenntnisse, Intelligenz, Lern- und Gedächtnisstrategien, Lernmotivation, Anstrengungsbereitschaft, Ausdauer, Selbstvertrauen und schulisches Selbstkonzept. Sie stehen in wechselseitigem Zusammenhang mit den strukturellen Merkmalen einer Familie, d. h., das Lernpotential eines Lerners und die Merkmale seiner Familie bedingen sich gegenseitig.

Lehrperson. Die Lehrkraft stellt einen zentralen Aspekt im Angebots-Nutzungs-Modell dar. Nach Helmke nimmt sie auf fünf Ebenen Einfluss auf das Unterrichtsangebot: 1. durch das Professionswissen (z. B. fachwissenschaftliche und fachdidaktische Expertise, schulpädagogisches und unterrichtsmethodisches Wissen, schultheoretische und bildungssoziologische Kenntnisse); 2. durch die fachliche, didaktische und diagnostische Kompetenz im konkreten Umgang mit der Klasse; 3. durch individuelle pädagogische Orientierungen und Überzeugungen (z. B. schul- und unterrichtsrelevante Werte, Kompetenzorientierung im Unterricht); 4. durch persönliche Erwartungen und Ziele (z. B. Leistungserwartungen, Einhaltung des Curriculums); sowie 5. durch Engagement, Geduld und Humor.

Unterricht. In Helmkes Modell stellt der Unterricht das Angebot dar, das den Schülerinnen und Schülern von der Lehrperson dargeboten wird und das diese annehmen, in Lernaktivitäten nutzen und für sich in einen Ertrag umwandeln können. Das Unterrichtsangebot umfasst sowohl die fachübergreifende als auch die fachspezifische Prozessqualität des Unterrichts, die sich beispielsweise aus dem adäquaten Einsatz von Sozialformen und fachdidaktischen Unterrichtsmethoden ergibt, sowie die Qualität des Lehr- und Lernmaterials. Ein weiterer Aspekt des Unterrichtsangebots stellt die Menge an nominal gewährter, tatsächlich unterrichteter und aktiv genutzter Unterrichtszeit dar. Weiterhin hängen Art und Umfang des Ertrags auch von der Wahrnehmung und Interpretation des Unterrichtsangebots durch die Lerner ab.

Lernaktivitäten. Wie bereits beim Aspekt „Unterricht“ erwähnt, führt die alleinige Darbietung des Unterrichtsangebots nicht zu einem Ertrag für die Lerner, z. B. hinsichtlich ihrer Kompetenzen. Das Unterrichtsangebot muss vielmehr mithilfe von Lernaktivitäten seitens der Schülerinnen und Schüler genutzt werden, um eine Wirkung zu erzielen. Diese Aktivitäten laufen im Unterricht (in einem zeitlich begrenzten Rahmen und mehr oder weniger stark von der Lehrkraft gesteuert) ab, erstrecken sich aber auch auf außerschulische Lernangebote. Die Quantität und Qualität der Lernaktivitäten wird darüber hinaus von Mediationsprozessen bestimmt, die den Lernprozess äußerst individuell werden lassen. Denn im Zuge der Nutzung des Unterrichtsangebots laufen verschiedenste Verarbeitungsprozesse ab, wie z. B. die persönliche Wahrnehmung und Interpretation des Unterrichtsangebots durch die Schülerinnen und Schüler sowie allgemeine Lern- und Denkprozesse, Motivation und Emotionen der Lerner.

Diese sechs Aspekte des Angebots-Nutzungs-Modells, die die Wirkungsweise des Unterrichts nach Helmke ausmachen, können – in Abhängigkeit von obig erwähnten Faktoren – zu einer Wirkung bzw. zu einem Ertrag seitens der Schülerinnen und Schüler führen. In Anlehnung an den produktorientierten Ansatz (Helmke, 2012: 21f.), besteht dieser nach Helmke in fachlichen und fachübergreifenden Kompetenzen sowie in der erzieherischen Wirkung der Schule. Neben der Entwicklung solcher Kompetenzen gibt es jedoch auch andere Zielkriterien, die als „Ertrag“ angestrebt werden könnten, wie beispielsweise Motivationsentwicklung oder die Egalisierung von Leistungsunterschieden innerhalb einer Klasse (vgl. Abschnitt 7.2).

1.1.2 Das Angebots-Nutzungs-Modell im Kontext des bilingualen Unterrichts

Aufgrund der vielfältigen, das Unterrichtsgeschehen beeinflussenden Aspekte, die im Angebots-Nutzungs-Modell berücksichtigt werden, eignet sich dieses für das Forschungsvorhaben der vorliegenden Arbeit. Denn bilingualer Unterricht soll möglichst umfassend und interdisziplinär analysiert werden, um adäquate Aussagen zu seiner Wirksamkeit und spezifischen Wirkungsweise treffen zu können. Hierfür steht insbesondere der Ertrag bilingualen Unterrichts für die Schülerinnen und Schüler im Mittelpunkt (vgl. Abschnitt 1.5), der als Maß für die Bestimmung der Wirksamkeit dieser Unterrichtsform sowohl bezüglich der Fremdsprachen- als auch der Sachfachkenntnisse dient.

Die Qualität des Unterrichts stellt hierbei einen wichtigen, diesen Ertrag beeinflussenden Faktor dar. Denn neben allgemeinen Qualitätsmerkmalen, die für jeden Unterricht zutreffen, weist der bilinguale Unterricht methodisch-didaktische Vorgehensweisen auf, die sich von denen des regulär einsprachig gehaltenen Unterrichts unterscheiden und somit als Spezifikum in diesem Kontext von besonderem Interesse sind (vgl. Abschnitt 1.4). Darüber hinaus ist auch die individuelle Wahrnehmung des Unterrichtsangebots, beispielsweise in Form von Emotionen der Schülerinnen und Schüler, bedeutsam. Diese werden domänen- bzw. kontextspezifisch erlebt (Götz, Cronjaeger, Frenzel, Lüdtke & Hall, 2010) und haben somit eine eigene, spezifische Wirkung auf den Ertrag seitens der bilingual unterrichteten Schülerinnen und Schüler (vgl. Abschnitt 1.4). Weiterhin sind auch die Lehrkräfte hinsichtlich der Wirksamkeit bilingualen Unterrichts von besonderem Interesse, da diese durch ihre Person die Qualität dieses Unterrichts mitbestimmen und somit die Schülerleistungen beeinflussen. Einzelne Überlegungen verweisen darauf, dass dies zugunsten der bilingual unterrichteten Klassen ausfallen könnte, da deren Lehrkräfte eine positiv selektierte Gruppe im Vergleich zu ihren einsprachig unterrichtenden Kolleginnen und Kollegen zu sein scheinen, sich bezüglich Schule und Unterricht also besonders motiviert und interessiert zeigen (Dalton-Puffer et al., 2010: 282; vgl. Abschnitt 1.3). Abschließend trägt ein weiterer zentraler Aspekt zum Schülerertrag bei, nämlich die Lernvoraussetzungen der Schülerinnen und Schüler bzw. Klassen, die durch Familie und Lernpotential bestimmt werden. Da die bilingual unterrichtete Schülerschaft eine spezifisch selektierte Lernergruppe zu sein scheint (Bruton, 2011), sind auch diese Faktoren von speziellem Interesse für die vorliegende Arbeit (vgl. Abschnitt 1.2). Andere Kontextfaktoren hingegen, die sich nicht auf die Zusammensetzung der Klasse beziehen, scheinen für Untersuchungen zur Wirksamkeit bilingualen Unterrichts nur begrenzt relevant. Dies hängt zu großen Teilen vom Studiendesign ab. Setzt sich beispielsweise eine Stichprobe aus Schülerinnen und Schülern aus Deutschland und aus nur einer Schulform zusammen, können Faktoren wie kulturelle Rahmenbedingungen und Schulform im Sinne statistischer Kontrollvariablen vernachlässigt werden. Dasselbe trifft auf den didaktischen Kontext zu, wenn sich der Forschungsgegenstand auf nur einen didaktischen Kontext im Sinne einer speziellen Unterrichtsform (der bilinguale Unterricht im Sachfach Geschichte) bezieht.

Diese Übertragung des Angebots-Nutzungs-Modells auf den Kontext des bilingualen Unterrichts macht deutlich, dass es für entsprechende Wirksamkeitsanalysen einer interdisziplinären Zusammenarbeit zwischen verschiedenen Wissenschaftsdomänen bedarf. Helmkes (2012) Modell weist Aspekte auf, die in die spezifischen Zuständigkeitsbereiche verschiedener Fachdisziplinen fallen. So sind etwa die Kategorien „Familie“ und „Lernpotential“ bzw. „Klassenzusammensetzung“ der empirischen Bildungsforschung bzw. Pädagogischen Psychologie zuzuordnen, die sich klassischerweise mit Personen- und Familienmerkmalen in verschiedenen Bildungskontexten beschäftigen (Reinders, Ditton, Gräsel & Gniewosz, 2010) und in diesem Zusammenhang über besondere Expertise im Bereich der Erfassung von kognitiven Fähigkeiten, Lernmotivation oder des

sozioökonomischen Hintergrunds von Familien verfügen. Die Analyse des „Unterrichtsangebots“ hingegen, mit dessen fachspezifischen methodisch-didaktischen Aspekten, fällt in den Zuständigkeitsbereich der Fachdidaktik, während allgemeine Merkmale der Unterrichtsqualität sowie die Wahrnehmung des Unterrichtsangebots in Form von Emotionen entweder in der empirischen Bildungsforschung oder in der Schulpädagogik verortet werden können (Reinders et al., 2010; Meyer, 2005). Dasselbe trifft auf die „Lehrperson“ zu. In Abhängigkeit vom bestehenden Forschungsinteresse kann dieser Aspekt im Hinblick auf pädagogische Orientierungen oder Engagement und Geduld der Lehrkraft der empirischen Bildungsforschung und/oder der Schulpädagogik zugeordnet werden (Kunter, Baumert, Blum, Klusmann, Krauss & Neubrand, 2011; Meyer, 2005). Liegt der Fokus auf Professionswissen sowie fachlichen und didaktischen Kompetenzen, fällt die „Lehrperson“ in den Verantwortungsbereich der Fachdidaktik. Schließlich lässt sich auch die „Wirkung“ des Unterrichts in der Fachdidaktik verorten, wenn als Zielkriterium die fachspezifischen Kompetenzen der Schülerinnen und Schüler definiert werden. Eine Besonderheit im bilingualen Unterricht stellt in diesem Zusammenhang die Tatsache dar, dass sich der Ertrag sowohl auf die sprachlichen als auch auf die sachfachlichen Fertigkeiten der Lerner bezieht, es also der Expertise der Fremdsprachen- sowie der Sachfachdidaktik bedarf. Zusammenfassend scheint der Verzicht auf eine Integration dieser verschiedenen Wissenschaftsdomänen in Forschungsarbeiten entweder zu einer unvollständigen oder zumindest unzureichend „fachfremd“ durchgeführten Analyse der Wirksamkeit bilingualen Unterrichts zu führen.

1.1.3 Zusammenfassung und Diskussion

Zusammenfassend eignet sich Helmkes (2012) Angebots-Nutzungs-Modell mit seinen vielfältigen Aspekten der Wirkungsweise des Unterrichts gut für die Übertragung auf den Kontext des bilingualen Sachfachunterrichts, da es verschiedenste Anknüpfungspunkte bietet, die für Analysen zur Wirksamkeit bilingualen Unterrichts von besonderem Interesse sind. Auf Grundlage der Überlegungen im vorhergehenden Abschnitt (1.1.2) sind diese 1. der familiäre Hintergrund, 2. das Lernpotential der Schülerinnen und Schüler, 3. die daraus resultierenden Lernvoraussetzungen auf Klassenebene, 4. die Lehrperson, 5. das Unterrichtsangebot und 6. dessen Wahrnehmung sowie 7. der Ertrag, der sich unter Berücksichtigung dieser Merkmale ergibt. Obige Überlegungen zeigten jedoch auch, dass es einer interdisziplinären Zusammenarbeit zwischen der empirischen Bildungsforschung sowie der Fremdsprachen- und Sachfachdidaktik bedarf, um eine angemessene Erfassung des bilingualen Unterrichts auf Basis des Angebots-Nutzungs-Modells zu gewährleisten.

Dennoch muss einschränkend angemerkt werden, dass auch Daten, die mithilfe einer umfassenden und interdisziplinären Erhebung gewonnen wurden, lediglich Annäherungen an die Unterrichtswirklichkeit darstellen können. Denn das Angebots-Nutzungs-Modell basiert auf mehreren theoretischen Annahmen, die dieses noch komplexer werden lassen. Erstens besteht eine

unbegrenzte wechselseitige Kompensierbarkeit der verschiedenen Aspekte und ein additives Zusammenwirken ist möglich. So sind beispielsweise die zahlreichen Ursachen von Schülerleistung teilweise miteinander verknüpft und können ausgeglichen oder ersetzt werden (Prinzip der Kompensation und Substitution). Zweitens bestehen sowohl wechselseitige Zusammenhänge zwischen Umwelt (Unterricht) und Individuum (Schüler/in) als auch hinsichtlich verschiedener Merkmale innerhalb eines Individuums. So beeinflussen beispielsweise die Vorkenntnisse eines Lernalters nicht nur dessen Ertrag (Kompetenzzuwachs), sondern sind gleichzeitig selbst abhängig von anderen Aspekten, wie Motivation oder Bildungsnähe der Eltern (vgl. Helmke, 2012: 28). Befunde zur Wirksamkeit bilingualen Unterrichts, die auf dem Angebots-Nutzungs-Modells basieren, müssen somit immer unter Berücksichtigung dieser Einschränkungen interpretiert werden.

Abbildung 1 zeigt zusammenfassend auf, welche Aspekte des Angebots-Nutzungs-Modells für Analysen zur Wirksamkeit bilingualen Unterrichts als relevant identifiziert wurden und somit in der vorliegenden Arbeit Berücksichtigung fanden. Dementsprechend werden nun in den folgenden Abschnitten 1.2 bis 1.5 die Kategorien „Familie“, „Lernpotential“ und „Kontext“, „Lehrperson“, „Unterricht“ und dessen „Wahrnehmung“ sowie „Wirkung“ nochmals konkret auf den Kontext des deutsch-englischen bilingualen Geschichtsunterrichts bezogen und für ihre Verwendung in den empirischen Teilstudien inhaltlich näher definiert.

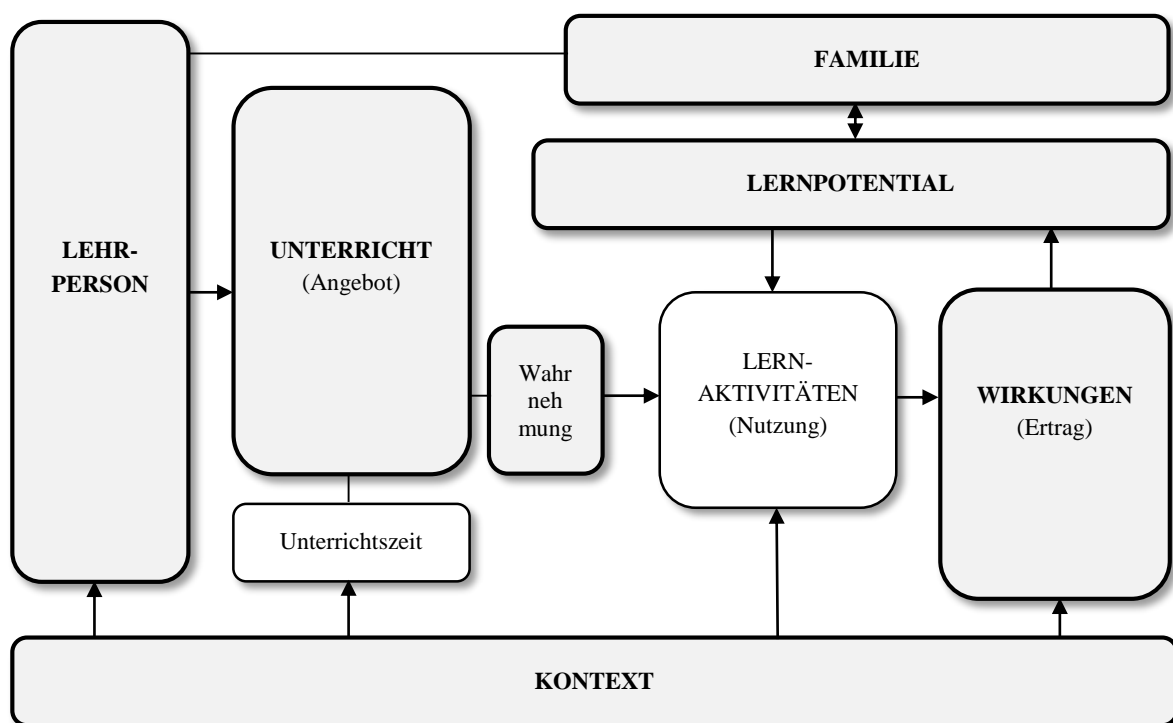


Abb. 1: Darstellung der in der vorliegenden Arbeit berücksichtigten Aspekte (grau hinterlegt) des Angebots-Nutzungs-Modells (mod. nach Helmke, 2012)

1.2 Die Schülerklientel des bilingualen Unterrichts

1.2.1 Organisation und Schülerschaft bilingualer Unterrichtsprogramme

Um die Schülerklientel des bilingualen Unterrichts im Sinne der Kategorien „Familie“ und „Lernpotential“ des Angebots-Nutzungs-Modells darstellen zu können, muss zunächst die Organisation bilingualer Unterrichtsangebote beschrieben werden. Der bilinguale Unterricht in Deutschland findet überwiegend innerhalb zweier Organisationsformen statt: Schulen mit fest etabliertem bilingualen Zug und Schulen, die bilinguale Module anbieten. Beide Organisationsformen finden sich inzwischen in allen Bundesländern bezüglich verschiedener Schulformen und in fast allen Schulfächern wieder (KMK, 2013: 11-14). Schulen mit bilingualen Modulen bieten bilingualen Unterricht nur phasenweise an, also nicht durchgängig durch das ganze Schuljahr. Die Schülerinnen und Schüler solcher Schulen nehmen im Verlauf eines Jahres ein oder mehrere Male an einigen bilingual unterrichteten Schulstunden oder an einer ganzen bilingualen Unterrichtseinheit („Module“) teil. In welchem Umfang jedoch, in welchen Fächern und in welchen Klassen(-stufen) bilinguale Module erteilt werden, ist der Schule freigestellt und hängt von den personalen Ressourcen der Schulen sowie von den individuellen Vorlieben der einzelnen bilingual unterrichtenden Lehrkräfte ab. Das bilinguale Unterrichtsangebot solcher Modul-Schulen ist somit nicht auf von vornherein fest ausgewählte Lernergruppen beschränkt, sondern kann prinzipiell von jeder Klasse – unter Berücksichtigung eben genannter Faktoren – wahrgenommen werden. In diesem Kontext kann demzufolge nicht von einer speziellen Schülerschaft gesprochen werden, die an bilingualen Unterrichtsangeboten teilnimmt.

Demgegenüber stehen von den Kultusministerien zertifizierte Schulen mit fest etabliertem bilinguaem Zug. Diese bieten lediglich einer ausgewählten Gruppe von Schülerinnen und Schülern durchgängig bilingualen Unterricht an. Durchgängig bedeutet, dass bilingualer Unterricht meist ab Klassenstufe 5 bis zum Ende der Schulzeit in mindestens einem Fach pro Schuljahr erteilt wird. Die Fächer variieren von Klassenstufe zu Klassenstufe. Deren Abfolge wird in den meisten Bundesländern durch das entsprechende Ministerium vorgegeben (z. B. KM, 2008: 48; KM, 2012). Weiterhin wird in manchen Bundesländern und Schulformen, beispielsweise an Gymnasien in Baden-Württemberg, dem bilingual unterrichteten Sachfach eine extra Unterrichtsstunde pro Woche zugewiesen, wodurch sich die nominale Unterrichtszeit im betroffenen Schulfach für bilingual Unterrichtete im Vergleich zu einsprachig Unterrichteten erhöht. Außerdem sehen viele bilinguale Zug-Schulen bilingualen Unterricht erst ab Klasse 7 vor. In diesem Fall wird für gewöhnlich in den Klassenstufen 5 und 6 vorbereitender Fremdsprachenunterricht vorgeschaltet, d. h., die Schülerinnen und Schüler nehmen in der Regel an zwei zusätzlichen Englischstunden pro Woche teil. Möchte ein Lerner an einem solchen regelmäßigen bilingualen Unterrichtsangebot teilnehmen, muss er sich im Anschluss an die Grundschule, vor Beginn der 5. Klasse, für einen bilingualen Zug einer solchen

Schule anmelden. D. h., Schülerinnen und Schüler eines bilingualen Zugs haben sich bewusst für diese Spezialisierung und somit gegen andere Profile ihrer Schule entschieden, wie etwa einen naturwissenschaftlichen Zug oder eine Musik- oder Sportklasse. Die Lerner des bilingualen Zugs bilden in der Regel eine gemeinsame Klasse und bleiben im Normalfall bis zum Ende der Schulzeit in diesem Zug. In manchen Zug-Schulen kommen die bilingual unterrichteten Lerner einer Klassenstufe jedoch nur für den bilingualen Unterricht zusammen und werden in allen anderen Fächern (inklusive Englischunterricht) in ihren Stammklassen unterrichtet. Unabhängig von der Organisationsform ist der Wechsel in andere Profile meist nur in Ausnahmefällen gestattet.

Im Gegensatz zu Schulen mit bilingualen Modulen bildet sich bei Zug-Schulen, deren Lerner sich aktiv für das bilinguale Unterrichtsangebot entscheiden, eine spezielle Schülerschaft heraus. Denn da meist mehr Nachfrage nach bilingualem Unterricht besteht als Plätze in den Klassen zur Verfügung stehen, finden Auswahlprozesse an den Schulen statt (Hollm, Hüttermann, Keßler & Schlemminger, 2010: 164). In der Regel gibt es zunächst Informationsveranstaltungen, in denen die verantwortlichen Lehrkräfte die besonderen Herausforderungen hervorheben, die der bilinguale Unterricht für Schülerinnen und Schüler mit sich bringt, wie etwa höhere kognitive Anforderungen, mehr Hausaufgaben, mehr Schulstunden pro Woche und die Notwendigkeit einer ausgeprägten Lernbereitschaft (Rumlich, 2013: 184; Bredenbröker, 2000: 50). Interessierte Eltern geben der Schule danach eine Bewerbung für ihr Kind ab, woran sich ein Beratungs- und Auswahlgespräch zwischen Eltern, Kind und Schulvertretung anschließt. Im Folgenden wählt die Schulleitung auf Basis der Bewerbung, des Gesprächs und des Grundschulzeugnisses passende Schülerinnen und Schüler aus. Falls es zu viele Kinder gibt, kann auch ein Losverfahren eingesetzt werden (Zydati, 2007: 70). Dieses Prozedere gibt Anlass zur Annahme, dass vor allem besonders motivierte und interessierte Kinder mit hoher Lernbereitschaft in bilinguale Klassen aufgenommen werden, die bereits in der Grundschule überdurchschnittlich erfolgreich waren, und deren eher bildungsnahe Eltern möglicherweise auch engagierter und interessierter die Schullaufbahn ihrer Kinder verfolgen. Zusammenfassend scheint also, dass Schülerinnen und Schüler bilingualer Züge eine spezielle Schülerschaft darstellen, die aufgrund der Selektivität der schulischen Auswahlverfahren über günstigere Lernvoraussetzungen verfügen könnten als die anderen Lerner ihrer Schule, wie beispielsweise über höhere kognitive Grundfähigkeiten, stärkeres schulisches Engagement, größere Motivation und ein vorteilhafteres Elternhaus. Somit kann erwartet werden, dass sich diese Vorteile auf alle schulischen Bereiche auswirken und zu besseren Fähigkeiten in den verschiedensten Schulfächern führen. Darüber hinaus ist es durch den zusätzlichen Fremdsprachenunterricht in den Klassenstufen 5 und 6, der an vielen Zug-Schulen als vorbereitende Maßnahme erteilt wird (z. B. KM, 2008), sehr wahrscheinlich, dass sich bereits zu diesem Zeitpunkt, vor Beginn des eigentlichen bilingualen Unterrichts, weitere Vorteile insbesondere im Bereich fremdsprachlicher Kompetenzen und Motivation herausbilden.

1.2.2 Familie und Lernpotential bilingual unterrichteter Schülerinnen und Schüler

Angesichts dieser Überlegungen zur möglichen Selektivität bilingualer Unterrichtsprogramme an bilingualen Zug-Schulen (vgl. Abschnitt 1.2.1) ist es für Untersuchungen zur Wirksamkeit dieser Unterrichtsform von besonderem Interesse, die Schüler- und die damit in Zusammenhang stehenden Familienmerkmale bilingual unterrichteter Lerner näher zu betrachten, um potentielle Unterschiede zur regulär einsprachig unterrichteten Schülerschaft feststellen zu können. Hierfür eignet sich das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts, da sich in ihm all diejenigen Aspekte wiederfinden, hinsichtlich derer im vorherigen Abschnitt (1.2.1) mögliche Vorteile zugunsten der bilingual unterrichteten Lerner angenommen wurden. Diese Aspekte sind Schichtzugehörigkeit und Bildungsnähe des Elternhauses sowie Intelligenz, Motivation und Vorkenntnisse der Lerner. Diese schüler- und familienbezogenen Merkmale werden im Folgenden hinsichtlich ihrer Verwendung in der vorliegenden Arbeit inhaltlich präzisiert sowie deren Relevanz im Kontext des bilingualen Unterrichts nochmals verdeutlicht.

Schichtzugehörigkeit. Die Schichtzugehörigkeit macht Aussagen über den sozialen Status eines Gesellschaftsmitglieds, wobei allen Mitgliedern einer Schicht ein vergleichbares soziales Prestige zugemessen wird (Springer Gabler, 2014). Die Zugehörigkeit zu einer sozialen Schicht kann mithilfe verschiedenster Kriterien bestimmt werden. Der strukturfunktionalistische Ansatz, der oft in fortgeschrittenen Industriegesellschaften vorzufinden ist, definiert Schichtzugehörigkeit insbesondere über die Merkmale Beruf, Bildung und Einkommen (Abels, 2009: 289). In Anlehnung an diese Definition wurde in der vorliegenden Arbeit der sozioökonomische Status zur Abbildung der Schichtzugehörigkeit der Schülerinnen und Schüler gewählt. Er beinhaltet Informationen zum Bildungsabschluss, der beruflichen Ausbildung und dem aktuellen Beruf der Eltern (vgl. Ganzeboom, de Graaf, Treiman & de Leeuw, 1992). Wie bereits erläutert, wird aufgrund der möglichen Selektivität bilingualer Unterrichtsprogramme in Zug-Schulen ein höherer sozioökonomischer Status zugunsten der bilingual unterrichteten Schülerinnen und Schüler angenommen. Dasselbe trifft auf die Bildungsnähe der entsprechenden Familien der Lerner zu.

Bildungsnähe. Der Begriff „Bildungsnähe“ verbreitete sich zu Beginn des neuen Jahrtausends insbesondere durch die Ergebnisse der PISA-Studie, wonach „bildungsferne“ Kinder schlechter in der Schule abschnitten als Kinder von Akademikern (PISA-Konsortium, 2001). Die Bildungsnähe einer Familie bezieht sich vorwiegend darauf, wie gebildet die Mitglieder dieses Personenkreises aufgrund ihrer schulischen und beruflichen Ausbildung sowie ihres alltäglichen Umgangs mit bildungsförderlichen Medien sind (ibid.: Kap. 8). Hieran anknüpfend wird Bildungsnähe in der vorliegenden Arbeit als die Gesamtheit mehrerer, der Bildung zuträglicher Aspekte verstanden: Bildungsabschluss und berufliche Ausbildung der Eltern, Anzahl der Bücher und das Vorhandensein einer Tageszeitung im Elternhaus.

Intelligenz. Intelligenz bezeichnet eine sehr allgemeine kognitive Fähigkeit, die ein Individuum beispielsweise dazu befähigt, abstrakt, schlussfolgernd und problemlösend zu denken (Gottfredson, 1997: 13). Nach Cattell (1963) lässt sich die kristalline Intelligenz (bereits erworbenes Lebenswissen und prozedurale Strategien) von der fluiden Intelligenz (Fähigkeit zur Aufnahme und Verarbeitung neuer Informationen) unterscheiden. Intelligenz kann sich aber auch auf verschiedene Fähigkeitsbereiche beziehen. Thurstone (1938) definierte diesbezüglich neun mentale Primärfähigkeiten: *spatial-visual abilities* (räumlich-visuelles Denken), *perceptual abilities* (Wahrnehmungsfähigkeit bzw. Auffassungsgabe), *numerical abilities* (rechnerisch-mathematische Fähigkeiten), *verbal relations* (Sprachverständnis), *word fluency* (Wortflüssigkeit), *memory* (Gedächtnisleistung), *derivation of rules* (Regelbildungsfähigkeit), *reasoning* (logisches Schlussfolgern) und *deduction* (Deduktionsfähigkeit). In Anlehnung an diese Definition wurden in der vorliegenden Arbeit zwei unterschiedliche Fähigkeitsbereiche berücksichtigt: die verbalen und figuralen kognitiven Grundfähigkeiten der Schülerinnen und Schüler (vgl. Heller & Perleth, 2000). Denn abgesehen von der Annahme, dass bilingual unterrichtete Lerner grundsätzlich über günstigere Lernvoraussetzungen und somit auch über eine höhere Intelligenz verfügen könnten als ihre Mitschülerinnen und Mitschüler (vgl. Abschnitt 1.2.1), ist angesichts der häufigen Verwendung der Fremdsprache in dieser Unterrichtsform zu erwarten, dass die Lerner insbesondere im verbalen Bereich Vorteile aufweisen (Fehling, 2008: 191). Die figuralen Fähigkeiten wurden kontrastiv hierzu gewählt.

Motivation. Motivation stellt die angenommene Ursache einer Aktivität dar (Wagner, 2009: 48) und bezieht sich somit auf einen „psychische[n] Prozess, der die Initiierung, Steuerung, Aufrechterhaltung und Evaluation zielgerichteten Handelns leistet“ (Dresel & Lämmle, 2011: 81). Für den schulischen Bereich ist insbesondere die Leistungsmotivation als Subdimension der Lernmotivation von Bedeutung (Wild, Hofer & Pekrun, 2001: 220-222). Das Aufbringen von Leistungsmotivation zur Durchführung einer Handlung hängt, dem Erwartungs-Mal-Wert-Modell von Eccles (z. B. Wigfield & Eccles, 2000) zufolge, von drei Faktoren ab: von der Erwartbarkeit eines Handlungserfolgs, vom subjektiven Wert des Handlungsergebnisses und gegebenenfalls von den Folgen des Ergebnisses. Hieran anknüpfend wird in der vorliegenden Arbeit Motivation als ein Zusammenspiel zwischen einem Erwartungsaspekt (akademisches Selbstkonzept der Lerner) und vierer Wertkomponenten (intrinsischer Wert, Wichtigkeitswert, Nützlichkeit, Kosten) verstanden. Für den Kontext des bilingualen Unterrichts wird erwartet, dass die Schülerinnen und Schüler aufgrund der selektiven Auswahlverfahren an bilingualen Zug-Schulen (vgl. Abschnitt 1.2.1) über eine grundsätzlich stärkere Leistungsmotivation verfügen als die einsprachig unterrichteten Lerner – und zwar sowohl in der Fremdsprache als auch im Sachfach.

Vorkenntnisse. Im schulisch-unterrichtlichen Kontext beziehen sich Vorkenntnisse in der Regel auf bereits vorhandenes Wissen der Schülerinnen und Schüler in verschiedenen Unterrichtsfächern. Darunter fallen sowohl inhaltliche Kenntnisse als auch fachspezifische

Strategien und Fertigkeiten („Kompetenzen“; vgl. Bach & Timm, 2013: 260), um sich diese Inhalte anzueignen und anzuwenden. Im Kontext des bilingualen Unterrichts sind, aufgrund dessen spezifisch dualer Ausrichtung, die Vorkenntnisse von mindestens zwei Domänen zu berücksichtigen: fremdsprachliche Fähigkeiten der Lerner sowie bereits vorhandenes Wissen und Kompetenzen im bilingual unterrichteten Sachfach. Weitere Ausführungen dazu, welche konkreten Fähigkeiten in der vorliegenden Arbeit in den beiden Fächern berücksichtigt wurden und warum, finden sich in Abschnitt 1.5, der sich mit Schülerleistungen im bilingualen Unterricht beschäftigt. Aus diesem Grund wird an dieser Stelle auf weitere Darstellungen verzichtet.

1.2.3 Bilingual unterrichtete Klassen

Die Überlegungen des vorhergehenden Abschnitts 1.2.2 verwiesen darauf, dass sich Vorteile zugunsten der bilingual unterrichteten Lerner bezüglich des sozioökonomischen Status, der Bildungsnähe des Elternhauses, der kognitiven Grundfähigkeiten, der Leistungsmotivation und der Vorkenntnisse ergeben könnten. Der Grund hierfür könnten mögliche Selektionsprozesse darstellen, die im Zusammenhang mit den Auswahlverfahren an bilingualen Zug-Schulen stattzufinden scheinen (vgl. Abschnitt 1.2.1). Hieran anknüpfend besteht weiterhin Grund zur Annahme, dass sich solche Unterschiede nicht nur auf bilingual und einsprachig unterrichtete Schülerinnen und Schüler beziehen, sondern sich vielmehr auf ganze bilingual und einsprachig unterrichtete Klassen erstrecken. Diese Argumentation bezieht sich auf das „Kontext“-Merkmal „Klassenzusammensetzung“ in Helmkes (2012) Angebots-Nutzungs-Modell, welches die Komposition einer Klasse in Bezug auf verschiedenste Aspekte wie Vorkenntnisniveau oder Schichtzusammensetzung in den Blick nimmt. Ausgehend von der eben berichteten Annahme zu Vorteilen bilingual Unterrichteter wäre also zu schlussfolgern, dass sich diese Vorteile nach einer arithmetischen Mittelung der individuellen Schülerwerte auch auf Klassenebene zeigen. Es scheint also wahrscheinlich, dass Unterschiede in den Lernvoraussetzungen nicht nur zwischen den bilingual und einsprachig unterrichteten Schülerinnen und Schülern, sondern als Konsequenz dessen auch zwischen ganzen bilingual und einsprachig unterrichteten Klassen bestehen. Diese würden sich zugunsten der bilingual unterrichteten Schülergruppen beispielsweise in einem durchschnittlich höheren sozioökonomischen Status, einer durchschnittlich größeren Bildungsnähe des Elternhauses sowie in im Klassendurchschnitt besseren kognitiven Grundfähigkeiten, einer höheren Motivation und größeren Vorkenntnissen manifestieren.

1.2.4 Zusammenfassung und Diskussion

Die vorhergehenden Abschnitte konnten zeigen, dass bilingual unterrichtete Schülerinnen und Schüler aus Schulen mit bilinguaalem Zug, aufgrund der dort stattfindenden Auswahlprozesse, sehr wahrscheinlich eine positiv selektierte Gruppe darstellen, die über günstigere Lernvoraussetzungen

verfügt als regulär einsprachig unterrichtete Lerner. In diesem Zusammenhang konnten in Anlehnung an das Angebots-Nutzungs-Modell fünf Aspekte aus den Bereichen „Familie“ und „Lernpotential“ identifiziert werden, bezüglich derer Grund zur Annahme besteht, dass sich Unterschiede zwischen bilingual und einsprachig unterrichteten Schülerinnen und Schülern zeigen. Es werden Vorteile zugunsten der bilingual unterrichteten Lerner bezüglich des sozioökonomischen Status und der Bildungsnähe des Elternhauses angenommen sowie höhere (verbale) kognitive Grundfähigkeiten, eine größere Leistungsmotivation und mehr Vorkenntnisse sowohl in der Fremdsprache als auch im Sachfach Geschichte. Aufgrund dessen werden – unter Bezug auf das Kontextmerkmal „Klassenzusammensetzung“ (Helmke, 2012) – dieselben Vorteile zugunsten bilingual Unterrichteter auch auf Klassenebene erwartet. Solche möglichen Unterschiede zwischen bilingual und einsprachig Unterrichteten sind von Interesse, wenn es darum geht, „die“ bilinguale Schülerschaft zu definieren und von anderen Lernergruppen abzugrenzen. Darüber hinaus stellt dies eine entscheidende Vorarbeit für Analysen zur Wirksamkeit bilingualen Unterrichts dar. Denn mithilfe der Informationen zu bereits bestehendem Vorwissen können Leistungsergebnisse beispielsweise daraufhin untersucht werden, welche Lernergruppe – bilingual oder einsprachig unterrichtete Schülerinnen und Schüler – über einen bestimmten Zeitraum hinweg mehr oder weniger Kompetenzen erworben hat. Gleichzeitig können hierbei unterschiedliche Ausgangslagen im Sinne günstigerer oder nachteiligerer Lernvoraussetzungen der Schülerinnen und Schüler bzw. Klassen berücksichtigt und statistisch konstant gehalten werden. Ein solches Vorgehen ist für Wirksamkeitsanalysen zu bilinguaem Unterricht wünschenswert, da diese Aspekte die Leistungsentwicklung der Lerner beeinflussen. Somit kann der tatsächliche Effekt dieser Unterrichtsform auf die Leistungsentwicklung bestmöglich isoliert werden.

1.3 Die bilingual unterrichtenden Lehrkräfte

1.3.1 Merkmale bilingual unterrichtender Lehrkräfte

Im vorhergehenden Abschnitt (1.2) wurde aufgezeigt, dass bilingual unterrichtete Lerner bzw. Klassen aus Schulen mit bilinguaem Zug eine spezifische Schülerklientel mit besonders günstigen Lernvoraussetzungen darstellen könnte. Auf ähnliche Weise bestehen mittlerweile erste Überlegungen dazu, dass sich auch die bilingual unterrichtenden Lehrkräfte in verschiedenen Merkmalen von ihren einsprachig unterrichtenden Kolleginnen und Kollegen unterscheiden könnten. Ein Ausgangspunkt dieser Annahme stellt die viel berichtete Tatsache dar, dass bilingual unterrichtende Lehrkräfte deutlich mehr Zeit in die Vorbereitung bilingualler Unterrichtsstunden investieren müssen, da nur wenige oder unpassende Materialien vorhanden sind und diese erst aufwendig erstellt oder überarbeitet werden müssen (Hollm, 2013; Kollenrott, 2008; Gierlinger, Hametner & Spann, 2007; Viebrock, 2007; Zydatis, 2007). Daran anknüpfend merkten Dalton-Puffer et al. (2010: 282) an, dass bilingual unterrichtende Lehrkräfte in der Hinsicht besonders seien,

dass sie gewillt wären, eine beträchtliche Menge an Mehrarbeit zu leisten. Den Autorinnen zufolge impliziert dies für gewöhnlich mehr Motivation und ein gesteigertes pädagogisches Interesse im Vergleich zu Lehrkräften im Allgemeinen. Abgesehen von diesem Aspekt des besonderen Engagements, der auch von Hollm (2013: 185) festgestellt wurde, kann darüber hinaus angenommen werden, dass bilingual unterrichtende Lehrerinnen und Lehrer möglicherweise auch ihre eigene berufliche Ausbildung erfolgreicher absolviert haben als ihre Kolleginnen und Kollegen und somit über bessere Voraussetzungen im Sinne von Professionswissen (vgl. Shulman, 1986) verfügen. Darauf deuten Überlegungen hin, wonach bilingual unterrichtende Lehrkräfte über überdurchschnittlich hohe Kompetenzen zu verfügen scheinen (Hollm, 2013: 185). Diese könnten beispielsweise von Auslandsaufenthalten im Zielsprachenland herrühren, die sich positiv auf die Fremdsprachenkompetenzen der Lehrerinnen und Lehrer auswirken können (Helmke & Klieme, 2008: 302; Küppers, 2008: 179f.; Müller-Hartmann & Schocker-von Ditfurth, 2008: 136). Dieser Aspekt scheint insbesondere für den Kontext des bilingualen Unterrichts interessant, da dieser überwiegend in der Fremdsprache gehalten wird (Diehr, 2012; vgl. Abschnitt 1.4.2) und somit gute Fremdsprachenfähigkeiten seitens der Lehrkräfte erforderlich macht.

1.3.2 Die bilingual unterrichtende Lehrperson im Rahmen des Angebots-Nutzungs-Modells

Obige Überlegungen (vgl. Abschnitt 1.3.1) geben Grund zur Annahme, dass bilingual unterrichtende Lehrkräfte möglicherweise über mehr schul- und unterrichtsbezogene Motivation und Engagement sowie über eine bessere berufliche Qualifikation verfügen als ihre einsprachig unterrichtenden Kolleginnen und Kollegen. Es ist somit von besonderem Interesse, dieses spezifisch mit bilinguaem Unterricht in Zusammenhang stehende Phänomen in den Blick zu nehmen, um mehr Informationen zu solchen potentiellen Unterschieden innerhalb der Lehrerschaft zu erhalten. Hierfür eignet sich das Angebots-Nutzungs-Modell, da es in der Kategorie „Lehrperson“ mehrere Anknüpfungspunkte für die für bilingual unterrichtende Lehrkräfte als spezifisch identifizierten Merkmale bietet. Diese sind das „Professionswissen“ bzw. die berufliche Expertise sowie schul- und unterrichtsbezogene Eigenschaften, Orientierungen und Einstellungen der Lehrkräfte, z. B. „Erwartungen und Ziele“ und „Engagement, Geduld und Humor“. Diese Aspekte werden im Folgenden in Bezug auf ihre Verwendung in der vorliegenden Arbeit konkretisiert.

Professionswissen. In Anlehnung an Shulman (1986) setzt sich lehrerbezogenes Professionswissen bzw. die Expertise einer Lehrkraft aus drei Arten des Lehrerwissens zusammen: Fachwissen, fachspezifisches pädagogisches (d. h. fachdidaktisches) Wissen und fachübergreifendes, allgemein pädagogisch-psychologisches Wissen (Helmke, 2012: 110). Das Angebots-Nutzungs-Modell knüpft an diese Definition an und berücksichtigt darüber hinaus auch Aspekte von Weinerts Ansatz zu Lehrerexpertise. Folglich gliedert sich Lehrerkompetenz in Helmkes (2012: 71) Modell in die folgenden vier Bereiche auf: Fachkompetenzen, didaktische

Kompetenzen, Klassenführungskompetenzen und diagnostische Kompetenzen. In der vorliegenden Arbeit wurde der Fokus auf den Aspekt der Fachkompetenz gelegt, da sich dieser – als Besonderheit des bilingualen Unterrichts – auf zwei Domänen erstreckt, nämlich auf die Fremdsprachenfähigkeiten der Lehrkräfte sowie auf deren Fachwissen im bilingual unterrichteten Sachfach. Für die Erfassung der Fremdsprachenfähigkeiten wurde exemplarisch der Aspekt des Auslandsaufenthalts in einem englischsprachigen Land gewählt – unter Rückbezug auf die Annahme, dass sich dieser günstig auf die Englischkenntnisse der Lehrkraft auswirken kann (vgl. Abschnitt 1.3.1). Für das Fachwissen in Geschichte wurde exemplarisch die Einsicht der Lehrkräfte in das Prinzip der Konstruktivität von Geschichte gewählt (vgl. Schreiber, Körber & van Borries, 2006). Zur Begründung und näheren Erläuterung der Auswahl dieser spezifischen Fähigkeit, die auch von Schülerinnen und Schülern eingefordert wird, finden sich ausführliche Darstellungen im Abschnitt 1.5.2, der sich mit den fachspezifischen Kompetenzen der Lerner im bilingualen Geschichtsunterricht beschäftigt. Schließlich wurde das Professionswissen der Lehrkräfte in der vorliegenden Arbeit weiterhin dadurch fokussiert, dass die Abschlussnoten der Lehrkräfte in ihren studierten Fächern berücksichtigt wurden. Auch diesbezüglich könnten bilingual unterrichtende Lehrkräfte, angesichts ihrer angenommenen höheren Qualifikation (vgl. Abschnitt 1.3.1), besser abschneiden als ihre einsprachig unterrichtenden Kolleginnen und Kollegen.

Erwartungen und Ziele. Schul- und unterrichtsrelevante Erwartungen und Ziele der Lehrperson können sich sowohl auf die Lerner (z. B. Leistungserwartungen, Intention und Wirkung von Belohnungen und Strafen) als auch auf die Lehrkraft an sich beziehen (z. B. berufsbezogene Selbstwirksamkeitserwartung, beruflicher Ehrgeiz; vgl. Schaarschmidt & Fischer, 2008). Da bisherige Überlegungen Grund zur Annahme geben, dass sich bilingual und einsprachig unterrichtende Lehrkräfte insbesondere hinsichtlich ihrer eigenen Einstellung zu Schule und Unterricht unterscheiden könnten (z. B. Motivation, Interesse; vgl. Abschnitt 1.3.1), wurden in der vorliegenden Arbeit diejenigen Erwartungen und Ziele in den Blick genommen, die die Lehrkraft an sich selbst stellt. In diesem Zusammenhang wurden folgende 10 arbeitsbezogene Verhaltens- und Erlebensmuster (Schaarschmidt & Fischer, 2008) gewählt, um diese Erwartungen und Ziele sowie weitere schul- und unterrichtsbezogene Eigenschaften und Einstellungen der Lehrkräfte abzubilden: Selbstwirksamkeitserwartung, beruflicher Ehrgeiz, Perfektionsstreben, subjektive Bedeutsamkeit der Arbeit, Verausgabungsbereitschaft, offensive Problembewältigung, innere Ruhe, Distanzierungsfähigkeit, emotionale Erschöpfung, Resignationstendenz.

Engagement, Geduld, Humor. Abschließend stellt insbesondere das Engagement einer Lehrperson für Schule und Unterricht einen Aspekt dar, für den für bilingual unterrichtende Lehrkräfte eine positivere Ausprägung angenommen wird als für ihre einsprachig unterrichtenden Kolleginnen und Kollegen (vgl. Abschnitt 1.3.1). Engagement bezeichnet einen meist auf Freiwilligkeit basierenden, intensiven Einsatz für eine Sache aus einer gefühlten Verbundenheit heraus (Duden, 2013). In dieser Hinsicht kann Engagement dem Perfektionsstreben (s. Absatz

Erwartungen und Ziele) zugeordnet werden. Darüber hinaus wurde es in der vorliegenden Arbeit über zwei weitere Aspekte berücksichtigt: Enthusiasmus für das Unterrichten (vgl. Helmke, 2012: 114) und Motivation (nach Wigfield & Eccles, 2000) für den bilingualen Unterricht.

1.3.3 Zusammenfassung und Diskussion

In den vorhergehenden Abschnitten wurde gezeigt, dass auch die bilingual unterrichtenden Lehrkräfte eine Gruppe mit speziellen Merkmalen darstellen könnte, bezüglich welcher sie sich möglicherweise von ihren einsprachig unterrichtenden Kolleginnen und Kollegen unterscheiden. Diese Annahme beruht auf ersten Überlegungen, wonach bilingual unterrichtende Lehrkräfte besonders engagiert und motiviert zu sein scheinen und über eine höhere berufliche Qualifikation verfügen könnten als andere Lehrpersonen (vgl. Abschnitt 1.3.1). In Anlehnung an das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts, wurden konkretisierend für die vorliegende Arbeit folgende Aspekte identifiziert, hinsichtlich derer solche Unterschiede zwischen bilingual und einsprachig unterrichtenden Lehrkräften erwartet werden können: 1. Fachkompetenzen als Teil des Professionswissens in Englisch und Geschichte (Auslandsaufenthalt in einem englischsprachigen Land, Einsicht in das Prinzip der Konstruktivität von Geschichte, Abschlussnoten im Studium); 2. verschiedene schul- und unterrichtsbezogene Erwartungen, Ziele und Einstellungen der Lehrkräfte (Selbstwirksamkeitserwartung, beruflicher Ehrgeiz, Perfektionsstreben, subjektive Bedeutsamkeit der Arbeit, Verausgabungsbereitschaft, offensive Problembewältigung, innere Ruhe, Distanzierungsfähigkeit, emotionale Erschöpfung, Resignationstendenz); sowie 3. Enthusiasmus und Motivation für das bilinguale Unterrichten als Ausdruck des Lehrerengagements. Mögliche Unterschiede in diesen Aspekten sind zunächst einmal per se von Interesse, da sie Erkenntnisse über „die“ bilingual unterrichtende Lehrerschaft liefern und somit Vergleiche zu anderen (z. B. einsprachig unterrichtenden) Lehrergruppen ermöglichen. Weiterhin ist die Berücksichtigung solcher möglicherweise positiveren Attribute seitens der bilingual Unterrichtenden wichtig, da Lehrkräfte durch ihre Person und ihren Unterricht Schülerleistung grundsätzlich beeinflussen (Helmke, 2012; Hattie, 2009; Klieme & Rakoczy, 2008) und somit zu Vorteilen für bilingual unterrichtete Klassen führen können. Deshalb sollten Unterschiede zwischen bilingual und einsprachig unterrichtenden Lehrpersonen in Untersuchungen zur Wirksamkeit bilingualen Unterrichts statistisch konstant gehalten werden, um den tatsächlichen Effekt dieser Unterrichtsform auf die Schülerkompetenzen bestmöglich von solchen verzerrenden Aspekten bereinigen zu können.

1.4 Das bilinguale Unterrichtsangebot

1.4.1 Methodisch-didaktische Aspekte im bilingualen Unterricht

Aus theoretischer Perspektive wird das Unterrichtsangebot jeden Unterrichts durch fachspezifische Vorgaben in Form didaktisch-methodischer Prinzipien bestimmt. Somit besitzt jedes Schulfach eine einzigartige Ausrichtung, die die Grundlage des Lehren und Lernens in diesem Fach darstellt. Dies gilt selbstverständlich auch für den bilingualen Unterricht, der spezifische methodisch-didaktische Merkmale aufweist, die ihn von anderem Unterricht unterscheiden und die deshalb für den vorliegenden Kontext von besonderem Interesse sind. Zu diesen Fachspezifika wurden zwar bereits erste Annäherungen geleistet (z. B. Thürmann, 2010; Breidbach, 2007), dennoch besteht bisher noch keine einheitliche und allgemein anerkannte Didaktik des bilingualen Sachfachunterrichts (Thürmann, 2013: 230; Rautenhaus, 2010: 109; Breidbach, 2010: 174). Aktuell tendiert die fachdidaktische Diskussion zur Etablierung sachfachspezifischer Didaktiken (vgl. Hallet & Königs, 2013), womit separate methodisch-didaktische Grundlagen für jedes bilingual unterrichtete Sachfach zur Verfügung ständen. Trotz dieses Trends bestehen vielzählige Aspekte, bezüglich derer im Rahmen des bilingualen Unterrichts noch keine vollständige theoretisch und/oder empirisch basierte Etablierung stattgefunden hat. Hierzu zählen beispielsweise (fehlende) methodisch-didaktische Leitlinien zu instruktionsbezogenem vs. konstruktivistischem Lehrstil, zum Einsatz von Sozialformen, zu Interdisziplinarität, zur Reduktion der Inhalte auf das Wesentliche und einem entsprechenden Einsatz von Materialien und (neuer) Medien, zu Unterstützungssystemen bei Verständnisschwierigkeiten, zur Korrektur von Schülerfehlern bzw. Schülerfeedback, zur Bewertung von Schülerleistung sowie zu Lern- und Arbeitstechniken.

Weiterhin stellt auch die Rolle der Sprachen im bilingualen Unterricht ein Aspekt dar, bezüglich dessen noch keine allgemein gültigen methodisch-didaktischen Grundprinzipien vorliegen (Heimes, 2013: 351; Keßler & Schlemminger, 2013; Diehr, 2012: 29). Hierbei steht die Frage der tatsächlichen Zweisprachigkeit dieser Unterrichtsform im Mittelpunkt, also Art und Umfang des Einsatzes von Mutter- und Fremdsprache. Denn obwohl der Terminus „bilingual“ auf die Präsenz zweier Sprachen hindeutet, scheint bilingualer Unterricht bisher eher als eine einsprachig in der Fremdsprache gehaltene Unterrichtsform verstanden zu werden. Dies bildet sich beispielsweise in den Bildungs- und Lehrplänen der Bundesländer Bremen, Hessen, Niedersachsen und Rheinland-Pfalz ab (vgl. Deutscher Bildungsserver, 1996-2014), in denen explizit erwähnt wird, dass bilingualer Unterricht Sachfachunterricht in der Fremdsprache sei und deshalb auch nur die Fremdsprache „konsequent und ausschließlich“ (Bildungsserver Rheinland-Pfalz, 1996-2014: 7) verwendet werden solle. Auch das europäische Informationsnetzwerk zu nationalen Bildungssystemen Eurydice verweist darauf, dass dies eine Unterrichtsform sei „in which a second language is used to teach certain subjects in the curriculum“ (Eurydice, 2006: 8). Die Muttersprache bleibt unerwähnt. Darüber

hinaus scheint bilingualer Unterricht auch in der Schulpraxis bisher überwiegend einsprachig in der Fremdsprache abzulaufen (Diehr, 2012: 18; Böing & Palmen, 2012: 73), was ebenfalls Befürworter in der aktuellen fachdidaktischen Debatte findet. So wies beispielsweise Heine (2010: 211) darauf hin, dass die gesteigerte Verarbeitungstiefe von Inhalten, die von ihr bei bilingual unterrichteten Schülerinnen und Schülern festgestellt wurde, gerade durch den umfassenden Einsatz der Fremdsprache möglich sei.

Neben dieser in Bezug auf den Einsatz der Sprachen einseitigen Auslegung des bilingualen Unterrichts bestehen jedoch auch Positionen, die den Einsatz der Muttersprache im Sinne eines „Verständnis-Notnagels“ sowie für spezifisches Fachvokabular tolerieren. So räumte Bonnet (2007: 130) ein, dass sich die Muttersprache im Zusammenhang mit Fachbegriffen tatsächlich als hilfreich erweisen könne. Theis (2010: 52) bestätigte, dass der Einsatz der Muttersprache eigentlich nicht erforderlich sei, jedoch einen überschaubaren Zugewinn im Bereich der Fachtermini erbringen möge. Eine solche Haltung wird auch von den Bildungs- und Lehrplänen in Hamburg, Sachsen und Thüringen vertreten (Deutscher Bildungsserver, 1996-2014). Weitere ähnliche Konkretisierungen finden sich in Broschüren zu bilinguaem Unterricht vom Ministerium für Kultus, Jugend und Sport Baden-Württemberg, in denen diese Unterrichtsform mit „weitgehender Einsprachigkeit“ bei gleichzeitiger „terminologischer Zweisprachigkeit“ (KM, 2006: 9) in Verbindung gebracht und darauf hingewiesen wurde, dass die Erreichung kognitiv anspruchsvoller Lernziele „einen Rückgriff auf die deutsche Sprache erforderlich machen“ könne (KM, 2008: 6). Abschließend besteht eine dritte Haltung zum Einsatz der Sprachen im bilingualen Unterricht, die die umfassende, reflektierte und systematische Mitbenutzung der Muttersprache vorsieht (vgl. Abschnitt 1.5.1: Butzkamms Prinzip der aufgeklärten Einsprachigkeit, 1973). Der Einsatz der Muttersprache nimmt demnach keine bloße „Notnagel-Funktion“ ein, sondern durch ihre Verwendung werden Vorteile im Lernprozess der Schülerinnen und Schüler durch systematische Unterstützung der Bereiche Verständnis, Ausdruck und Fachterminologie erwartet (Diehr, 2012: 25). Otten und Wildhage (2007: 19) argumentierten beispielsweise, dass die Muttersprache bei der Begriffs- und Kontextbildung von Bedeutung sei, da sie eine zusätzliche Reflexionsebene eröffne. Geiss (2013) bestätigte, dass im bilingualen Geschichtsunterricht lernförderliche Vergleiche nicht gezielt generiert werden müssten, sondern die Schülerinnen und Schüler „von sich aus auf semantische Differenzen in der Fachterminologie aufmerksam“ würden (ibid.: 57). Denselben Vorteil bilingualen Geschichtsunterrichts stellte Kollenrott (2008: 52f.) aus dem Blickwinkel interkulturellen Lernens fest (z. B. semantische Diskrepanz zwischen „Völkerwanderung“ und „barbarian invasion“; weitere Beispiele in Heimes, 2013: 349). Gierlinger et al. (2007: 37) gingen ebenfalls davon aus, dass die Muttersprache zu einem besseren Verständnis sachfachlicher Konzepte führe und daher nicht negativ behaftet sein solle. Auch Heimes (2013: 349) bewertete den Einbezug der Muttersprache im bilingualen Unterricht als unverzichtbar.

1.4.2 Der Einsatz der Sprachen als Spezifikum des bilingualen Unterrichts

Die Rolle der Sprachen im bilingualen Unterricht scheint angesichts obiger Überlegungen (Abschnitt 1.4.1) nach wie vor weitestgehend ungeklärt zu sein. Dies stellt jedoch das zentrale Merkmal dar, das den bilingualen vom einsprachigen Sachfachunterricht unterscheidet. Die Frage nach dem Einsatz von Mutter- und Fremdsprache beschreibt also das Spezifikum bilingualen Unterrichts, womit dessen nähere Betrachtung besonders lohnenswert scheint. Hierfür eignet sich das Angebots-Nutzungs-Modell, bei dem der Einsatz der Sprachen in der Kategorie „Unterricht“ verortet werden kann und somit ein Bestandteil des Angebots ist, das den Schülerinnen und Schülern von der Lehrkraft gemacht wird und das es zu nutzen gilt, um eine entsprechende Wirkung zu erzielen. Dem Angebots-Nutzungs-Modell zufolge kann die Qualität des Unterrichtsangebots sowohl über fachspezifische als auch über fachübergreifende Merkmale bestimmt werden. Die für den bilingualen Unterricht spezifische Fragestellung zur Verwendung von Mutter- und Fremdsprache gehört somit erster Kategorie an.

Die Konkretisierung dieser sprachbezogenen Aspekte erfolgt für die vorliegende Arbeit in Anlehnung an die Überlegungen des vorhergehenden Abschnitts, wonach zunächst grundsätzlich zu bedenken ist, ob die Muttersprache überhaupt Einsatz im bilingualen Unterricht findet, in welchem Ausmaß dies der Fall ist und ob dies gleichermaßen auf die Lehrkraft sowie die Schülerinnen und Schüler zutrifft. Weiterhin soll berücksichtigt werden, für welche Zwecke die Muttersprache verwendet wird. Die aktuelle fachdidaktische Diskussion (vgl. Abschnitt 1.4.1) deutet hierzu an, dass dies insbesondere für das Verständnis unterstützende Maßnahmen, wie etwa Wiederholungen und Zusammenfassungen von Unterrichtsstoff oder Erklärungen komplexer Sachverhalte, sowie für Fachtermini der Fall ist. Bezüglich der Fachbegriffe soll weiterhin konkretisiert werden, ob diese in Anlehnung an das Prinzip der terminologischen Zweisprachigkeit (vgl. Abschnitt 1.4.1) in beiden Sprachen, lediglich in einer Sprache oder ohne einer spezifischen Sprachregel folgend unterrichtet werden.

1.4.3 Unterrichtsqualität im Kontext des bilingualen Unterrichts

Dem Angebots-Nutzungs-Modell zufolge kann Unterrichtsqualität nicht nur über fachspezifische Merkmale, wie z. B. der Einsatz der Sprachen im bilingualen Unterricht (vgl. Abschnitt 1.4.2), sondern auch anhand von fachübergreifenden Aspekten bestimmt werden. Solche Qualitätsmerkmale können in der Regel Aussagen zur Güte des Unterrichtsangebots in jedem Schulfach treffen, besitzen also fachübergreifende Gültigkeit. Hierfür stehen verschiedene Modelle zur Prozessqualität des Unterrichts zur Verfügung (Helmke, 2012; Kunter et al., 2011; Klieme & Rakoczy, 2008), die allgemeine, für alle Schulfächer relevanten Aspekte der Unterrichtsqualität in Betracht nehmen, wie beispielsweise die Motivierung der Lerner oder ein lernförderliches Klassenklima. Diese fachübergreifende Ausprägung der Unterrichtsqualität scheint auch im Kontext des bilingualen

Unterrichts von besonderem Interesse. Denn auf Basis der Überlegungen aus Abschnitt 1.3, die zeigten, dass bilingual unterrichtende Lehrkräfte grundsätzlich motivierter, engagierter und qualifizierter sein könnten als ihre einsprachig unterrichtenden Kolleginnen und Kollegen, ist davon auszugehen, dass sich diese positiven Attribute auch auf die Qualität ihres Unterrichts auswirken (Köller et al., 2012: 348) – ein Zusammenhang der unabhängig vom bilingualen Unterricht bereits des Öfteren formuliert sowie empirisch belegt wurde (z. B. Kunter, Klusmann, Baumert, Richter, Voss & Hachfeld, 2013; Helmke, 2012). Demnach böten bilingual unterrichtende Lehrkräfte ihren Schülerinnen und Schülern qualitativ besseren Unterricht an als einsprachig unterrichtende Lehrerinnen und Lehrer.

Eine Konkretisierung dieses fachübergreifenden Aspekts der Unterrichtsqualität im Kontext des bilingualen Unterrichts scheint deshalb lohnenswert. Hierfür stehen zahlreiche Klassifikationen zur Verfügung (Helmke, 2012; Borich, 2007; Meyer, 2005; Slavin, 1997). In der vorliegenden Arbeit wurde das Modell von Klieme und Rakoczy (2008) gewählt, da es – im Gegensatz zu anderen Taxonomien – inhaltlich verdichtete, theoretisch abgeleitete sowie empirisch gestützte Kriterien zur Bestimmung von Unterrichtsqualität aufweist. Das Modell bezieht sich auf die drei Konstitutionsbedingungen und Ziele von Unterricht „Disziplin“, „Motivation“ und „Leistung“ nach Diederich und Tenorth (1997). Disziplin kann, Klieme und Rakoczy (2008) zufolge, durch störungspräventives Verhalten seitens der Lehrkraft sowie durch Klarheit und Strukturiertheit in einem möglichst lehrerkontrollierten Unterricht erreicht werden (vgl. direkte Instruktion; Mayer, 2004). Motivation kann weiterhin durch die Förderung des positiven emotionalen Erlebens erhöht werden (vgl. Selbstbestimmungstheorie von Deci & Ryan, 1993), wie beispielsweise durch wertschätzende Beziehungen zwischen Lehrkraft und Schülerinnen und Schülern oder positive Rückmeldungen an die Lerner. Die Leistungen der Schülerinnen und Schüler können schließlich auf Basis eines konstruktivistischen Grundverständnisses verbessert werden, welches besagt, dass Selbsttätigkeit der Schülerinnen und Schüler zu Lernerfolg führt – allerdings nur dann, wenn hierbei eine kognitive Aktivierung der Lerner erfolgt (Mayer, 2004: 17). Anknüpfend an diese Überlegungen leiteten Klieme und Rakoczy (2008) drei Basisdimensionen guten Unterrichts ab, die unter anderem in der TIMSS-Video-Studie von 1999 empirisch verifiziert wurden: 1. strukturierte, klare und störungspräventive Klassenführung; 2. konstruktive Unterstützung der Schülerinnen und Schüler durch die Lehrkraft mithilfe eines unterstützenden, schülerorientierten Sozialklimas; sowie 3. kognitive Aktivierung der Lerner. Diese Merkmale wurden in der vorliegenden Studie zur Berücksichtigung der fachübergreifenden Prozessqualität von Unterricht im Kontext des bilingualen Lehren und Lernens gewählt.

1.4.4 Individuelle Wahrnehmung des bilingualen Unterrichtsangebots

Helmkes (2012) Angebots-Nutzungs-Modell geht davon aus, dass das Unterrichtsangebot der Lehrkraft durch die Lerner aktiv genutzt werden muss, um zu einem Ertrag zu führen. Diese Nutzung

des Unterrichtsangebots in Form von Lernaktivitäten seitens der Schülerinnen und Schüler findet jedoch individuell unterschiedlich statt und wird beeinflusst von der persönlichen Wahrnehmung und Interpretation des Unterrichts. Helmke zufolge wird diese Unterrichtswahrnehmung über verschiedene Lern- und Denkprozesse sowie Motivation und Emotionen der Lerner gesteuert. Im Kontext des bilingualen Unterrichts kann insbesondere bezüglich Emotionen angenommen werden, dass diese bei bilingual unterrichteten Schülerinnen und Schülern positiver ausgeprägt sein könnten als bei einsprachig unterrichteten Lernern. Denn die Überlegungen aus Abschnitt 1.2 geben Grund zur Annahme, dass die bilingual unterrichtete Schülerschaft, die grundsätzlich interessierter und motivierter im Unterricht zu sein scheint, diesen auch positiver erlebt. Diese Annahme wird durch Befunde gestützt, wonach Motivation und unterrichtsbezogene Emotionen in enger Wechselwirkung zueinander stehen (Pekrun, Götz, Frenzel, Barchfeld & Perry 2011; Götz et al., 2010). Für den vorliegenden Kontext scheint also die Beantwortung der Frage interessant, welche Emotionen Lerner bezüglich des bilingualen Unterrichts empfinden und ob sich diese Emotionen von denjenigen unterscheiden, die einsprachig unterrichtete Schülerinnen und Schüler bezüglich desselben Sachfachs im einsprachigen Unterrichtssetting erleben.

Deshalb wurden in der vorliegenden Arbeit die Leistungsemotionen von Schülerinnen und Schülern als Manifestation des individuellen Unterrichtserlebens gewählt. Leistungsemotionen sind Emotionen in Bezug auf leistungsbezogene Aktivitäten und deren Ergebnisse (Pekrun, 2006: 317). Sie treten sowohl als situative, momentane Zustände (*states*) als auch als relativ stabile Persönlichkeitseigenschaften (*traits*) auf (Frenzel & Stephens, 2011: 22f.). In Rückbezug auf obige Überlegungen, wonach die Motiviertheit und das möglicherweise besonders positive Unterrichtserleben bilingual unterrichteter Lerner als stabile, charakterisierende Merkmale dieser speziellen Schülerschaft verstanden werden, sind für die vorliegende Arbeit insbesondere Leistungsemotionen im Sinne überdauernder Persönlichkeitseigenschaften relevant. Deshalb wurde Pekruns (2006) Taxonomie gewählt, welche auf der von ihm entwickelten Kontroll-Wert-Theorie basiert, und Leistungsemotionen als *traits* unter anderem in Abhängigkeit davon systematisiert, wann diese von den Lernern erlebt wurden: vor, während oder nach dem Unterricht (*prospective outcome emotions, activity emotions, retrospective outcome emotions*). Folglich wurden für die vorliegende Arbeit fünf Leistungsemotionen ausgewählt, die diesen temporären Aspekt von Pekruns Klassifikation abdecken: Hoffnung, Freude, Angst, Ärger, Langeweile.

1.4.5 Zusammenfassung und Diskussion

Mithilfe von Helmkes Angebots-Nutzungs-Modell als theoretischer Grundlage konnte in den vorhergehenden Abschnitten gezeigt werden, dass die nähere Betrachtung des bilingualen Unterrichtsangebots besonders lohnenswert ist, da es sich in dreierlei Hinsicht von anderen Lehrformen unterscheiden könnte: bezüglich fachspezifischer und fachübergreifender Unterrichtsqualitätsmerkmale sowie der Wahrnehmung des Unterrichtsangebots. Aus

fachspezifischer Perspektive sind spezielle methodisch-didaktische Aspekte gemeint, die den bilingualen vom regulär einsprachig gehaltenen Unterricht unterscheiden, wie beispielsweise der Einsatz von Mutter- und Fremdsprache. Neben dessen deskriptiver Beschreibung ist darüber hinaus von besonderem Interesse, wie sich solche Maßnahmen auf die Leistungen der bilingual unterrichteten Schülerinnen und Schüler auswirken. Denn mit entsprechenden Erkenntnissen könnte das bilinguale Unterrichtsangebot noch effektiver gestaltet und ein Beitrag zu einer empirisch fundierten Didaktik und Methodik des bilingualen Unterrichts geleistet werden. Deshalb wurden in der vorliegenden Arbeit folgende sprachbezogenen Merkmale des bilingualen Unterrichts berücksichtigt: Häufigkeit des Einsatzes von Mutter- und Fremdsprache durch Lehrkraft und Lerner, zweckgebundene Verwendung der Muttersprache (für Wiederholungen und Zusammenfassungen, Erklärung komplexer Sachverhalte, Fachtermini) sowie Sprachregel für die Einführung von Fachbegriffen (beide Sprachen, nur in einer Sprache, ohne spezifische Sprachregel).

Bezüglich der fachübergreifenden Qualitätsmerkmale gaben obige Überlegungen (vgl. Abschnitt 1.4.3) Grund zur Annahme, dass sich das Unterrichtsangebot bilingual unterrichtender Lehrkräfte positiver gestalten könnte als das der einsprachig unterrichtenden Kolleginnen und Kollegen. Denn für die bilingual unterrichtenden Lehrkräfte werden Vorteile in den Bereichen Fachkompetenz, schul- und unterrichtsbezogene Ziele und Einstellungen sowie Engagement angenommen (vgl. Abschnitt 1.3), was sich positiv auf die allgemeine Qualität des bilingualen Unterrichts auswirken kann. Schließlich scheint wahrscheinlich, dass auch die Wahrnehmung des Unterrichtsangebots durch bilingual unterrichtete Schülerinnen und Schüler positiver ausfällt als bei einsprachig unterrichteten Lernern. Basis für diese Annahme stellen Überlegungen dar, wonach die bilingual unterrichtete Schülerschaft über eine grundsätzlich höhere Motivation und Lernbereitschaft verfügt (vgl. Abschnitt 1.2), was sich günstig auf ihre im Zusammenhang mit Unterricht erlebten Emotionen auswirken kann. Solche Erkenntnisse zu möglichen Unterschieden in der fachübergreifenden Qualität des Unterrichtsangebots und in dessen Wahrnehmung zwischen bilingualem und einsprachigem Unterrichtskontext sind für beschreibende Analysen des bilingualen Unterrichts wertvoll. Darüber hinaus sind diese Erkenntnisse insbesondere dann relevant, wenn die Wirksamkeit dieser Unterrichtsform in den Blick genommen und hierbei Einflussfaktoren auf die Schülerleistung ausgeschlossen werden sollen, die nicht per se in Verbindung mit den methodisch-didaktischen Prinzipien des bilingualen Unterrichts stehen, sondern sich lediglich auf dessen besondere Lehrer- und Schülerschaft zurückführen lassen. Daher wurden in der vorliegenden Arbeit folgende Aspekte zum Unterrichtsangebot berücksichtigt, die für eine bestmögliche Isolierung der Effekte bilingualen Unterrichts statistisch kontrolliert werden sollten: Klassenführung, konstruktive Unterstützung durch die Lehrkraft und kognitive Aktivierung der Lerner (vgl. Klieme & Rakoczy, 2008) als Merkmale fachübergreifender Unterrichtsqualität sowie die Leistungsemotionen Hoffnung, Freude, Angst, Ärger und Langeweile (vgl. Pekrun, 2006) als Indikatoren der Unterrichtswahrnehmung.

1.5 Schülerleistungen im bilingualen Unterricht

1.5.1 Theoretische Fundierung des bilingualen Unterrichts

Der bilinguale Sachfachunterricht in Deutschland nahm seinen Ursprung im Jahre 1969, als der erste deutsch-französische bilinguale Zug an einem Gymnasium in Baden-Württemberg als Reaktion auf den deutsch-französischen Kooperationsvertrag von 1963 eingeführt wurde (KMK, 2013: 7). Ziel war es, die in Folge des Zweiten Weltkriegs beschädigten nachbarschaftlichen Verhältnisse bei der nachwachsenden Generation durch „gegenseitige Vermittlung von Sprache und Kultur“ (ibid.: 3) zu verbessern. Abgesehen von diesem soziokulturellen Aspekt, der vorwiegend in den Anfangsjahren des bilingualen Unterrichts sowie für die an der Rheinschiene gelegenen Schulen von großer Bedeutung war, etablierten sich in der fachdidaktischen Diskussion theoriebasierte sowie lernpsychologisch begründete Annahmen, wonach bilingual unterrichtete Schülerinnen und Schüler aufgrund der „generell andere[n], bessere[n] Rahmenbedingungen“ (Wolff, 1997: 51f.) auch anderweitig von dieser „überlegenen“ (ibid.) Unterrichtsform profitieren könnten. Dies kann sich beispielsweise auf bessere Leistungen in der Fremdsprache sowie im bilingual unterrichteten Sachfach beziehen – und zwar unabhängig davon, in welcher Fremdsprache und welches Sachfach bilingual unterrichtet wird. Ein Beispiel für einen solchen theoretischen Bezugspunkt, mithilfe dessen Vorteile zugunsten bilingual unterrichteter Lerner argumentiert werden können, stellt der Einsatz der Sprachen im bilingualen Unterricht dar; den Aspekt also, der das Spezifikum dieser Unterrichtsform darstellt (vgl. Abschnitt 1.4.2). Daher scheint die nähere Betrachtung dieses Zusammenhangs in der vorliegenden Arbeit lohnenswert. Im Folgenden werden nun einige theoretische Ansätze vorgestellt, die Grund zur Annahme liefern, dass sich der Einsatz der Sprachen positiv auf den Schülerertrag im bilingualen Unterricht auswirken kann und die somit einen Teil der theoretischen Fundierung dieser Unterrichtsform darstellen. Die Annahmen rekurren auf Zweitspracherwerbstheorien sowie kognitionspsychologische Erkenntnisse und thematisieren sowohl die Verwendung der Fremdsprache als auch den Miteinbezug der Muttersprache (vgl. Abschnitt 1.4).

Die Rolle der Fremdsprache. Es bestehen verschiedene theoretische Ansätze, die Grund zur Annahme liefern, dass sich die Fremdsprachenkenntnisse bilingual unterrichteter Schülerinnen und Schüler aufgrund der Verwendung einer Fremdsprache im bilingualen Unterricht zur Vermittlung sachfachlicher Inhalte besonders positiv entwickeln könnten: die Input-Hypothese (Krashen, 1985), der *natural approach* (Krashen & Terrell, 2000) sowie der *interaction approach* (Gass & Mackey, 2007). Krashens Input-Hypothese (1985) ist ein Bestandteil seines Monitormodells (ibid.) und nimmt an, dass die Konfrontation eines Lerners mit einer ausreichend großen Menge an verständlichem fremdsprachlichem Input – der sich typischerweise auf einem etwas höheren als dem eigenen Niveau befindet – besonders förderlich für den Erwerb der Fremdsprache ist. Der Input kann beispielsweise

dadurch verständlicher gestaltet werden, dass die Lehrperson Gebrauch von reduzierter Sprache (*foreigner talk*) oder von zusätzlichem sprachlichem sowie nicht-sprachlichem Material macht. Diese Bedingungen sind im bilingualen Unterricht gegeben, da die Lerner im Vergleich zu einsprachig unterrichteten Schülerinnen und Schülern einer größeren Menge an fremdsprachlichem Input durch ihre Lehrkräfte ausgesetzt sind, der sich darüber hinaus typischerweise auf einem etwas höheren sprachlichen Niveau als dem der Lerner bewegt (Thürmann, 2010: 78f.). Zur Vereinfachung dieses Inputs wenden bilingual unterrichtende Lehrkräfte in der Regel *foreigner talk* an und greifen auf weitere Unterstützungssysteme zurück, wie beispielsweise Vokabellisten oder Bilder (Viebrock, 2007; Müller-Schneck, 2006).

Anknüpfend an die Input-Hypothese und weitere Aspekte von Krashens (1985) Monitormodell stützt auch der *natural approach* (Krashen & Terrell, 2000) die Annahme, dass bilingual unterrichtete Schülerinnen und Schüler besonders bezüglich ihrer Fremdsprachenkompetenzen von dieser Unterrichtsform profitieren können. Diesem unterrichtsmethodisch ausgerichteten Ansatz zufolge wird die Fremdsprache dann besonders effektiv erlernt, wenn dies unter natürlichen Bedingungen geschieht, die denjenigen ähneln, die beim Erwerb der Muttersprache vorherrschten. Krashen und Terrell (2000) konkretisierten hierfür unter anderem die folgenden Merkmale: Eine auf natürliche Kommunikation ausgerichtete Zielsetzung des Unterrichts, welche durch die Verwendung interessanter und möglichst authentischer Kommunikationsinhalte und Lernaktivitäten erreicht werden soll. Kein Unterricht, dessen zentrales Ziel die Bewusstmachung grammatikalischer Formen darstellt (vgl. die Unterscheidung zwischen „focus on form(S)“ und „focus on meaning“; Long & Robinson, 2005). Weiterhin soll kein Zwang zur Sprachproduktion bestehen und Fehler in der schülerseitigen Sprachproduktion sollen nur dann korrigiert werden, wenn diese die inhaltliche Verständlichkeit beeinträchtigen. Diese Aspekte sind charakteristisch für den bilingualen Unterricht, da es dort nicht vorrangig um einen perfekt akkuraten Gebrauch der Fremdsprache geht (Diehr, 2013: 212; vgl. „message before accuracy“ in Timm, 2013: 203f.), sondern die Sprache lediglich zur Vermittlung von Inhalten eingesetzt wird („focus on meaning“). Dementsprechend orientiert sich die Fehlerkorrektur der Lehrkraft im bilingualen Unterricht in der Regel an einem „gentle feedback“ und findet vorwiegend bei Einschränkungen in der Verständlichkeit statt (KM, 2006: 15f.). Da die zu behandelnden Inhalte durch die Sachfachcurricula festgelegt werden, führt dies zu einem hohen Maß an Authentizität bezüglich dieser Inhalte und der entsprechenden Lernaktivitäten, da diese um ihretwillen gelernt bzw. ausgeführt werden und nicht aus reinen Übungszwecken zum Erwerb der Fremdsprache herangezogen wurden. Weiterhin müssen Schülerinnen und Schüler im bilingualen Unterricht in der Regel nicht ausschließlich in der Fremdsprache kommunizieren, sondern dürfen *code-switching* (Haugen, 1956: 40; vgl. Königs, 2013: 176; Böing & Palmen, 2012: 81) in die Muttersprache betreiben.

Schließlich geht auch Gass und Mackeys (2007) *interaction approach* davon aus, dass der verstärkte Umgang mit der Fremdsprache deren Erwerb vorantreibt. Demnach ist nicht nur die Menge an Fremdspracheninput durch die Lehrkraft wichtig, sondern auch die Häufigkeit und Qualität, in der Lerner mit dieser Fremdsprache interagieren und Feedback dazu erhalten, indem sie diese produzieren und ihre Hypothesen zum korrekten Sprachgebrauch mithilfe von entsprechenden Lehrerrückmeldungen anpassen (Thürmann, 2010: 79; vgl. Output-Hypothese von Swain, 1985, und Interaktions-Hypothese von Long, 1996). Auch hierfür bietet der bilinguale Unterricht ausreichend Möglichkeiten, da er genug Raum zur Kommunikation in der Fremdsprache bereithält, und zwar sowohl für die Sprachproduktion seitens der Schülerinnen und Schüler als auch für korrigierendes Feedback seitens der Lehrkräfte (jedoch nur, wenn diese Phasen der Spracharbeit im bilingualen Unterricht auch zulassen und entsprechende Fehlerkorrekturen betreiben).

Bilingualität – Der Miteinbezug der Muttersprache. Butzkamms Prinzip der aufgeklärten Einsprachigkeit (1973) bzw. funktionalen Fremdsprachigkeit (2002) geht weiterhin davon aus, dass sich auch der reflektierte und systematische Miteinbezug der Muttersprache im Fremdsprachenunterricht positiv auf den effektiven und umfassenden Erwerb der Fremdsprache auswirkt. Butzkamm (2002: 15-50) argumentierte in Rückbezug auf die Bilingualismusforschung und psycholinguistische Erkenntnisse, dass jeder Mensch von seiner Muttersprache geprägt sei und bei Kontakt mit einer Fremdsprache automatisch Verknüpfungen zu ihr vornehmen würde. Er führte aus, dass Lerner bereits über muttersprachenbezogenes semantisch-situatives und syntaktisch-formales Vorwissen sowie über weitere sprachliche Fähigkeiten, wie Lesen oder Schreiben, verfügten und dieses Wissen nur noch auf die Fremdsprache übertragen und etwas anpassen müssten. Somit sei der Teil, der im Zusammenhang mit der Fremdsprache umgelernt werden müsse, „nur ein Bruchteil dessen, was neu zu lernen wäre, wenn man noch keine Sprache hätte“ (Butzkamm, 2002: 24). Deshalb sollte dieses Unterstützungssystem Butzkamm zufolge nicht ignoriert oder unterdrückt, sondern muttersprachliche Verknüpfungen sollten bewusst berücksichtigt und planvoll in den Unterricht integriert werden. Der bilinguale Sachfachunterricht bietet hierfür ideale Möglichkeiten, da sowohl Lehrkräfte als auch Schülerinnen und Schüler in verschiedenen Situationen auf die Muttersprache zurückgreifen (vgl. Abschnitte 1.4.1, 1.6.4). Auch Butzkamm (2002: 150) bezeichnete diese Unterrichtsform als „Idealfall“ für die Anwendung seines Prinzips der aufgeklärten Einsprachigkeit. Denn anders als im regulären Fremdsprachenunterricht besteht Butzkamm zufolge im bilingualen Unterricht kaum die Gefahr, dass Lerner Wörter oder Satzmuster aus Übungszwecken bedeutungslos und automatisiert aneinander ketten würden. Vielmehr fände besonders bedeutungsvolle und authentische Kommunikation statt, die dem Fremdspracherwerbsprozess zusätzlich zugutekomme. Folglich kann angenommen werden, dass sich die Mitverwendung der Muttersprache im bilingualen Unterricht förderlich auf die Fremdsprachenfähigkeiten der Schülerinnen und Schüler auswirken könnte.

Darüber hinaus sollten sich auch positive Effekte auf die Kompetenzen der bilingual unterrichteten Lerner im Sachfach ergeben. Denn es darf angenommen werden, dass das verbesserte Verstehen der Sprache im Unterricht sowie das Behandeln authentischer Themen zu einem gesteigerten Verständnis und größerer Akzeptanz der Kommunikationsinhalte führt. Diese Annahme wird unterstützt von der Schwellenhypothese (Cummins, 1979). Sie besagt, dass in einem bilingualen Unterrichtssetting ausreichend große Kenntnisse in der Mutter- und Fremdsprache jenseits der oberen Kompetenzschwelle zu „enhanced metalinguistic, academic, and cognitive functioning“ (Cummins, 2006: 182) führen kann. Bestehen also ausreichend hohe Sprachkompetenzen, kann das bilinguale Unterrichtssetting einen zusätzlichen „linguistic and cognitive growth“ (Cummins, 2006: 175) bei den Lernern bewirken, der alle schulischen Domänen, sowohl (Fremd-)Sprachen als auch Sachfächer, betrifft. Einen solchen möglichen kognitiven Zugewinn führen Adesope et al. (2010) auf die kognitiven Korrelate des Bilingualismus zurück. Denn in ihrer Metaanalyse zeigten sich Vorteile bilingualer Lerner gegenüber einsprachigen Kindern in den Bereichen metalinguistisches und metakognitives Bewusstsein, abstrakte und symbolische Repräsentationen, Aufmerksamkeitskontrolle (vgl. Bialystok, 1999) sowie Problemlösestrategien. Adesope et al. (2010) zufolge werden diese Vorteile durch den kontinuierlichen Wechsel zwischen den beiden Sprachen verursacht, weshalb sich bilinguale Kinder in der Regel sowohl sprachlich als auch kognitiv aufmerksamer und flexibler zeigen und Informationen nachhaltiger verarbeiten als Lerner eines einsprachigen Settings. Einschränkend muss jedoch beachtet werden, dass sich solche kognitiven Vorteile möglicherweise nur begrenzt bei den bilingual unterrichteten Schülerinnen und Schülern des vorliegenden Kontexts zeigen könnten. Denn diese stellen keine im eigentlichen Sinne „bilingualen“ Kinder dar, die bereits von Geburt an mit zwei Sprachen aufgewachsen sind (vgl. Tej & William, 2008), sondern erlernten die Fremdsprache in der Regel erst im Schulunterricht. Dennoch bieten der verstärkte fremdsprachliche Input und die Möglichkeiten zur eigenen Sprachproduktion im bilingualen Sachfachunterricht optimale Bedingungen für die Lerner, ein ausreichend hohes Fremdsprachenniveau zu erreichen, um somit sowohl gesteigerte fremdsprachliche als auch sachfachliche Lernfortschritte zu initiieren, die im regulär einsprachigen Unterrichtssetting vielleicht nicht möglich wären.

Diese Überlegungen zu Theorien und Prinzipien aus den Bereichen Zweitspracherwerb und Kognitionspsychologie geben Grund zur Annahme, dass sich Vorteile zugunsten der bilingual unterrichteten Schülerinnen und Schüler bezüglich ihrer Sprach- und Sachfachkompetenzen im Vergleich zu einsprachig unterrichteten Lernern ergeben könnten. Deshalb scheint es lohnenswert, die Schülerkompetenzen im spezifischen Kontext des bilingualen Unterrichts näher zu betrachten.

1.5.2 Fachspezifische Kompetenzen im deutsch-englischen Geschichtsunterricht

In Helmkes Angebots-Nutzungs-Modell lassen sich Schülerkompetenzen in der Kategorie „Ertrag“ verorten, zu dem sowohl fachliche und fachübergreifende Kompetenzen als auch die erzieherische

Wirkung der Schule zählen. Aufgrund der obigen Überlegungen (Abschnitt 1.5.1) werden nun im Folgenden die fachspezifischen Kompetenzen des für die vorliegende Arbeit bedeutsamen deutsch-englischen Geschichtsunterrichts konkretisiert.

Nach Weinert (2001: 27f.) sind Kompetenzen „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“ Klieme, Avena, Blum, Döbrich, Gruber et al. (2003) übernahmen diesen Definitionsansatz und fügten hinzu, dass Kompetenzen eine Disposition sei, „die eine Person befähigt, bestimmte Arten von Problemen erfolgreich zu lösen, also konkrete Anforderungssituationen eines bestimmten Typs zu bewältigen“ (ibid.: 72). Sie erweiterten Weinerts Definition somit um die Domänenspezifität von Kompetenzen und folgerten daraus, dass sich jede Operationalisierung einer Kompetenz auf eine solche „konkrete Anforderungssituation“ beziehen müsse und deshalb auch in den jeweiligen Schulfächern separat zu erfolgen habe (ibid.: 73-75). Da der bilinguale Unterricht zwei Domänen vereint, Kompetenzen in der Fremdsprache sowie im Sachfach, werden im Folgenden beide Bereiche in den Blick genommen. Zuvor soll jedoch darauf hingewiesen werden, dass Kompetenzen lediglich kognitive Konstrukte sind (Weskamp, 2013: 261) und deshalb nur leistungsbezogen erfasst und gemessen werden können (Klieme et al., 2003: 73). Kompetenzen äußern sich also immer in der Performanz, in der tatsächlich erbrachten Leistung, der Schülerinnen und Schüler (Heil, 2012: 17; ISB Bayern, 2006; vgl. Chomsky, 1972). In diesem Sinne beabsichtigt die vorliegende Arbeit zwar, Schülerkompetenzen zu untersuchen, erfasst wurden diese jedoch in Form von Leistungen. Aus Praktikabilitätsgründen werden beide Begriffe synonym verwendet (vgl. DESI-Konsortium, 2006).

Englischkompetenzen. Nach Weskamp (2013: 257) befähigen Kompetenzen „zur Übertragung von Wissen auf andere Situationen und Zusammenhänge. Für den Englischunterricht bedeutet dies beispielsweise, dass Vokabeln nicht nur erlernt und abgefragt, sondern kommunikativ angewandt werden können.“ Diese Definition bildet den in den letzten Jahren vorherrschenden fachdidaktischen Zugang zu Englischkompetenzen ab, nämlich die Entwicklung kommunikativer Kompetenzen (nach Piepho, 1974) als übergeordnetes Ziel des Englischunterrichts (Krumm, 2007: 118). Kompetenzorientierung im Unterricht gewann bereits seit dem Jahr 2000, in Folge der Ergebnisse der PISA-Studie, immer mehr an Gewicht und wurde als Paradigmenwechsel zur Output-Orientierung verstanden (Mayer & Köhler, 2009: 4). Ein Jahr später wurde der „Gemeinsame Europäische Referenzrahmen für Sprachen“ (GER; Europarat, 2001) veröffentlicht, der als Ziele für den Fremdsprachenunterricht Handlungsorientierung und kommunikative Kompetenzen der Lerner vorsah (ibid.: 21-27). Auch die von der Kultusministerkonferenz im Jahre 2003 beschlossenen „Bildungsstandards für die erste Fremdsprache“ übernahmen diese Ausrichtung auf funktionale kommunikative Kompetenzen (KMK, 2004: 8). Es wurden vier Teilbereiche unterschieden, die dabei helfen sollen, die kommunikative Kompetenz der Schülerinnen und Schüler im

Fremdsprachenunterricht auszubilden. 1. kommunikative Fertigkeiten (Hören, Lesen, Sprechen, Schreiben, Sprachmittlung); 2. Verfügung über die sprachlichen Mittel (Wortschatz, Grammatik, Aussprache, Orthographie); 3. interkulturelle Kompetenz (z. B. praktische Bewältigung interkultureller Begegnungssituationen); sowie 4. methodische Kompetenzen (z. B. Lernstrategien, Präsentieren und Mediennutzung). Diese Kompetenzbereiche finden sich so oder ähnlich – in verschiedenen Zusammenstellungen und Gewichtungen – auch in den einschlägigen fachdidaktischen Veröffentlichungen zum Englisch- bzw. Fremdsprachenunterricht wieder (z. B. Hallet & Königs, 2010; Haß, 2009; Timm, 2005). Darüber hinaus wurde der Ansatz kommunikativer Kompetenzen weiterhin durch einen aktuelleren Beschluss der Kultusministerkonferenz von 2011 zur Stärkung der Fremdsprachenkompetenzen bestätigt, der die schriftliche und mündliche Sprachproduktion sowie Sprachrezeption, interkulturelles Lernen und Sprachbewusstheit als Indikatoren kommunikativer Kompetenz in den Mittelpunkt des Fremdsprachenunterrichts rückte (KMK, 2011: 3-5).

Um diese Vielzahl an Teilkompetenzen möglichst umfassend berücksichtigen zu können, folgt die vorliegende Arbeit einem Ansatz aus dem Bereich der Testwissenschaft, wonach Sprache als ganzheitliches Konstrukt (Klein-Braley, 1995: 499) bzw. eindimensionales Phänomen verstanden wird und davon ausgeht, dass der Sprachfähigkeit ein „einheitlicher, genereller Faktor“ (Zydatiś, 2005: 75) zugrunde liegt. Auf Grundlage dieser *unitary competence hypothesis* (Oller & Hinofotis, 1980) wird somit der allgemeine Sprachstand (Vollmer, 2007: 367) bzw. die *general language proficiency* (Grotjahn, 2011: 131) von Lernern fokussiert. Dementsprechend wurde in der vorliegenden Arbeit der C-Test gewählt, um eine solche tiefer liegende latente Sprachkompetenz abzubilden. Es ist ein Instrument zur globalen, integrativen Sprachstandsmessung, das verschiedene linguistische Teilkompetenzen abdeckt, wie beispielsweise Orthographie, Lexik, Morphologie, Syntax und Semantik (Zydatiś, 2005: 76-78). Der C-Test basiert auf dem Prinzip der Reduzierung sprachlicher Redundanzen und verlangt von den Schülerinnen und Schülern die zweite Hälfte jedes zweiten Wortes durch Rekonstruktionsleistung zu ergänzen (Grotjahn, 2011: 132). Er hat einen hohen diagnostischen Wert, insbesondere für die Erfassung von Sprachstandsentwicklungen (Zydatiś, 2005: 77), und ist sehr objektiv und reliabel (Grotjahn, 2011: 131). C-Tests werden daher häufig in large-scale Studien eingesetzt (z. B. DENOCS, Rumlich, 2012; DESI, DESI-Konsortium, 2008; DEZIBEL, Zydatiś, 2007). Dennoch muss einschränkend berücksichtigt werden, dass dieses Messinstrument keinen kommunikativen Test darstellt und daher nicht die tatsächliche Sprachverwendung der Lerner abbilden kann. Aus diesem Grund sollten C-Tests mit einem Sprachfähigkeitstest kombiniert werden (Zydatiś, 2005: 76), der die Fremdsprachenfähigkeiten von Schülerinnen und Schülern bezüglich einzelner Teilkompetenzen erfasst. Diesem Verständnis liegt die *divisible competence hypothesis* (Oller & Hinofotis, 1980) zugrunde, die im Gegensatz zur *unitary competence hypothesis* davon ausgeht, dass Sprachfähigkeit ein multidimensionales Phänomen darstellt. Deshalb wurde in der vorliegenden Arbeit neben dem C-Test auch das

Hörverstehen als Teilkompetenz der englischen Sprachfähigkeit berücksichtigt, um diese möglichst umfassend abzubilden. Denn insbesondere bezüglich dieser rezeptiven Fertigkeit werden Vorteile zugunsten der bilingual unterrichteten Lerner erwartet, und zwar aufgrund des verstärkten Englischinputs, dem die Lernenden durch ihre Lehrkräfte ausgesetzt sind (Köller et al., 2012: 337; Thürmann, 2010: 78; Nold et al., 2008: 452-454).

Geschichtskompetenzen. Obwohl der weiter oben beschriebene Paradigmenwechsel hin zur Output-Orientierung das gesamte deutsche Schulsystem betrifft und somit für alle Altersstufen, Schulformen und Fächer gilt, wurden von der Kultusministerkonferenz bisher noch nicht für alle Schulfächer Bildungsstandards herausgegeben. Dies trifft auch auf das für den vorliegenden Kontext relevante Schulfach Geschichte zu. Nichtsdestotrotz wurden in einzelnen Bundesländern Konkretisierungen in den Bildungs- und Lehrplänen vorgenommen (Heil, 2012: 50-70; z. B. Baden-Württemberg: „Angestrebte Kompetenzen sind die Narrativität [...] sowie die Fähigkeit zu Kritik, Argumentation, Urteilsbildung“; KM, 2004: 216), die für Lehrkräfte eine – verbindliche – Orientierung darstellen sollen. Abgesehen von solchen eher vage bleibenden ministeriellen Vorgaben zu angestrebten Schülerkompetenzen im Geschichtsunterricht kursieren in der fachdidaktischen Diskussion seit mehreren Jahren verschiedenste Kompetenzmodelle. Diese wurden vielfach diskutiert und adaptiert, dennoch besteht bisher keine Einigung auf ein Modell, das von der Mehrheit der Fachdidaktiker und Schulpraktiker unterstützt würde (Barricelli, Gautschi & Körber, 2012: 211; Heil, 2012: 48). Eine Gemeinsamkeit der meisten Modelle scheint jedoch darin zu liegen, dass ihnen die beiden geschichtstheoretischen Dimensionen Geschichtsbewusstsein und Narrativität zugrunde liegen – wenn auch unter Gebrauch unterschiedlicher Begrifflichkeiten (Bertram, Wagner & Trautwein, 2014: 196; Barricelli et al., 2012: 212). Im Folgenden werden nun exemplarisch drei Kompetenzmodelle (Gautschi, 2009; FUER-Gruppe bzw. Schreiber et al., 2006; Pandel, 2005) vorgestellt, um einen Überblick über Kompetenzen im Geschichtsunterricht zu erhalten und daraus diejenigen abzuleiten, die für die vorliegende Arbeit gewählt wurden.

Pandels (2005) Zugang zu Geschichtskompetenzen stellt das „älteste vollständig ausformulierte“ Modell Deutschlands dar (Barricelli et al., 2012: 217). Ihm zufolge gilt es im Geschichtsunterricht vier Grundkompetenzen an die Lerner zu vermitteln: 1. Die Gattungskompetenz zur Unterscheidung und zum richtigen Umgang mit verschiedenen Textgattungen; 2. die Interpretationskompetenz, um dem erworbenen Wissen Sinn entnehmen und mögliche Ambivalenzen entdecken zu können; 3. die narrative Kompetenz, um historische (Er-) Kenntnisse sinnbildend verbinden und somit eine Geschichte herstellen zu können; sowie 4. die geschichtskulturelle Kompetenz für die geschichtskulturell angemessene Deutung historischer Begebenheiten und eine daraus ableitbare Orientierung im Sinne zukünftigen verantwortlichen Handelns. Gautschi (2009) hingegen versteht – in Anlehnung an Rüsens (2008) Überlegungen zum Geschichtsbewusstsein – die narrative Kompetenz als das Hauptziel historischen Lernens. Dieses kann über die Ausbildung von vier Teilkompetenzen erreicht werden: 1. Wahrnehmungskompetenz,

um Veränderungen in der Zeit zu erkennen (z. B. zur Herleitung historischer Fragen); 2. Erschließungskompetenz bezüglich historischer Quellen und Darstellungen (zum korrekten Umgang mit verschiedenen Gattungen; befähigt zur Sachanalyse [vgl. Jeismann, 1980]); 3. Interpretationskompetenz für Geschichte (zur Herleitung von Sachurteilen); sowie 4. Orientierungskompetenz für Zeiterfahrungen (zur Sinnbildung und Orientierung im eigenen Leben; befähigt zu Werturteilen). Schließlich stellt das „Kompetenz-Strukturmodell historischen Denkens“ der FUER-Gruppe (Schreiber et al., 2006) den bisher am umfassendsten dargestellten und diskutierten Zugang zu Geschichtskompetenzen dar (Barricelli et al., 2012: 219). In Anlehnung an das Prozessmodell historischen Denkens (vgl. Hasberg & Körber, 2003 bzw. Rösen, 1983) identifizierte die Forschergruppe des Projekts „FUER Geschichtsbewusstsein“ vier Kompetenzbereiche, die in ihrem Prozessmodell verortet sind und die für eine Stabilität und Sicherheit sorgende Orientierung in der Gegenwart unter Zuhilfenahme geschichtlicher (Er-) Kenntnisse notwendig sind: 1. historische Fragekompetenz zur Initiation historischer Denk- und Orientierungsprozesse, indem Fragen an die Vergangenheit gestellt oder zugrunde liegende Fragestellungen von bereits bestehenden Narrationen erschlossen werden; 2. historische Methodenkompetenz zur fachgerechten Beantwortung historischer Fragen unter Zuhilfenahme von Quellen und Darstellungen (Re-Konstruktionskompetenz) sowie zur Analyse und Überprüfung der historischen Triftigkeit von bereits vorliegenden Darstellungen (De-Konstruktionskompetenz); 3. historische Orientierungskompetenz, um historische (Er-)Kenntnisse, also beispielsweise selbst rekonstruierte historische Narrationen oder eigene Stellungnahmen gegenüber einer historischen Darstellung, auf die Gegenwart und die eigene (zukünftige) Lebenssituation übertragen und wirksam machen zu können; sowie 4. historische Sachkompetenz, die durch das wiederholte Durchlaufen von Prozessen des historischen Denkens (Kompetenzbereiche 1. bis 3.) erworben, verändert und erweitert wird und sich auf die Fähigkeit bezieht, Geschichte mithilfe historischer Prinzipien, Konzepte und Kategorien sowie adäquaten fachspezifischen Begriffen erschließen, strukturieren und reflektieren zu können.

In der vorliegenden Arbeit wird zur Konkretisierung von Geschichtskompetenzen lediglich ein Teilbereich adressiert, und zwar die Einsicht der Schülerinnen und Schüler in das geschichtsbezogene epistemologische Prinzip der (Retro-)Perspektivität, Partikularität und Konstruktivität (Baumgartner, 1997), welches im FUER-Modell der historischen Sachkompetenz zugeordnet wurde (Bertram et al., 2014: 197). Hierbei geht es um die Einsicht, dass Geschichte eine abstrakte und von Menschen retrospektiv erzeugte Rekonstruktion darstellt und es daher notwendig ist „vorgefertigte“ historische Darstellungen zu dekonstruieren. Dies manifestiert sich im Schulalltag beispielsweise darin, dass sich die Schülerinnen und Schüler dem Unterschied zwischen den Gattungen Quelle und Darstellung sowie den Implikationen dieser Unterscheidung für den Umgang mit der konstruierten Geschichte bewusst sein sollten. Denn Darstellungen müssen als bereits vorliegende historische Narrationen in ihren Bestandteilen erfasst und auf ihre Tiefenstruktur

analysiert werden (De-Konstruktion). Quellen hingegen können zur Erschließung historischer Begebenheiten dienen, was durch die Entwicklung eigener Narrationen möglich wird (Synthese, Re-Konstruktion; Schreiber et al., 2006: 22-24). Die Fokussierung der Einsicht in dieses epistemologische Prinzip der Konstruktivität von Geschichte scheint insbesondere im Kontext des bilingualen Unterrichts von besonderem Interesse, da durch den Gebrauch der „Fremd“-Sprache Verfremdungseffekte erzeugt werden und somit eine Distanz zum Lerngegenstand entstehen kann (Hollm & Pirner, 2008: 117). Ein solch gewonnener Abstand ermöglicht es den Schülerinnen und Schülern möglicherweise, leichter zu erkennen, dass historische Texte kritisch betrachtet und hinterfragt, Narrationen also adäquat de- und rekonstruiert werden müssen, als dies beim Umgang mit vertrauten und glaubwürdiger erscheinenden muttersprachlichen Geschichtserzeugnissen der Fall wäre. Dies gäbe Grund zur Annahme, dass bilingual unterrichtete Schülerinnen und Schüler diese spezifische, auf das epistemologische Prinzip der Konstruktivität von Geschichte bezogene, Sachkompetenz besser ausbilden könnten als einsprachig unterrichtete Lerner (vgl. Heimes, 2013: 347).

Darüber hinaus wurde die Fähigkeit, den Konstruktionscharakter von Geschichte zu erkennen auch deshalb für die vorliegende Arbeit gewählt, da hierfür bereits ein validiertes „Kurzinstrument“ für den Einsatz in large-scale Studien vorlag (Bertram et al., 2014). Dies ist beispielsweise weder für die Erfassung der De- und Re-Konstruktionskompetenz des FUER-Modells (Schreiber et al., 2006), noch für die Gattungs- und Erschließungskompetenz nach Pandel (2005) bzw. Gautschi (2009) der Fall, welche auf den tatsächlichen (korrekten) Umgang der Schülerinnen und Schüler mit Quellen und Darstellungen abzielen – und somit nicht auf das bloße Erkennen des Konstruktionscharakters von Geschichte begrenzt bleiben. Angesichts dessen, dass solche in größerem Rahmen einsetzbare Forschungsinstrumente zu Geschichtskompetenzen zurzeit praktisch nicht vorhanden sind (Bertram et al., 2014: 196; eine Ausnahme hierzu bildet Hartmann, 2008), muss diese teilweise forschungsmethodisch bedingte Auswahl in der vorliegenden Arbeit hingenommen werden.

Abgesehen von der Einsicht in das Prinzip der Konstruktivität von Geschichte, soll in der vorliegenden Arbeit – unabhängig von den Kompetenzmodellen – das Sachwissen der Schülerinnen und Schüler in Geschichte berücksichtigt werden. „Wissen“ wird Klieme et al. (2003: 73) zufolge als Basiselement zur Ausbildung von Kompetenzen verstanden, um schließlich zu „Können“ (kompetentem Handeln) seitens der Lerner zu führen. Trotz der weiter oben beschriebenen Kompetenzorientierung in deutschen Klassenzimmern müssen Schülerinnen und Schüler also nach wie vor über konkretes sachfachliches Inhaltswissen verfügen, anhand dessen Kompetenzen exemplarisch eingeübt werden können. Für den Kontext des bilingualen Unterrichts scheint dieses Sachwissen, insbesondere hinsichtlich von Fachbegriffen und ihrer zugrunde liegenden Konzepte, von speziellem Interesse zu sein. Denn dem Fachvokabular kommt in dieser Unterrichtsform eine besondere Bedeutung zu, da es im Sinne terminologischer Zweisprachigkeit (vgl. Abschnitt 1.4.1)

sowohl in der Mutter- als auch in der Fremdsprache gelehrt werden sollte (Heimes, 2013: 349; Theis, 2010: 52; KM, 2006: 9). Dies soll, mitunter aufgrund von semantischen Diskrepanzen, zu tieferem Verständnis seitens der Lerner führen (vgl. Abschnitt 1.4.1.) und gibt somit Grund zur Annahme, dass bilingual unterrichtete Schülerinnen und Schüler im Bereich des Sachwissens über Vorteile gegenüber den einsprachig unterrichteten Lernern verfügen könnten, die Fachbegriffe und deren zugrunde liegenden Konzepte lediglich in einer Sprache dargeboten bekommen. Aufgrund dieser speziell im Kontext des bilingualen Unterrichts entstehenden Fragestellung wird das Geschichtswissen der Schülerinnen und Schüler in der vorliegenden Arbeit in Anlehnung an den relevanten Bildungsplan von Baden-Württemberg (KM, 2004) berücksichtigt und in Form von Fragen zu wichtigen historischen Personen, Daten und Begriffen sowie zu historischen Zusammenhängen, Gründen für und Konsequenzen von historischen Ereignissen konkretisiert.

1.5.3 Zusammenfassung und Diskussion

Die Übertragung verschiedener Theorien und Ansätze aus dem Bereich des Zweitspracherwerbs sowie kognitionspsychologischer Erkenntnisse auf den Kontext des bilingualen Unterrichts konnte zeigen, dass Schülerinnen und Schüler dieser Unterrichtsform bessere Leistungen als einsprachig unterrichtete Lerner zeigen könnten. Denn es wird angenommen, dass das Mehr an Fremdspracheninput sowie das interaktionsreiche Erlernen dieser Fremdsprache in einer natürlichen Spracherwerbsumgebung – Bedingungen, die der bilinguale Unterricht bietet – zu Vorteilen zugunsten seiner Schülerinnen und Schüler hinsichtlich der Entwicklung der Fremdsprachenkompetenzen führen können (Gass & Mackey, 2007; Krashen & Terrell, 2000; Krashen, 1985). Weiterhin können auch Unterschiede hinsichtlich der sachfachlichen Leistungen angenommen werden. Denn aufgrund des additiven Bilingualismus (Cummins, 1979), der durch einen umfassenden und systematischen Miteinbezug von Mutter- und Fremdsprache im bilingualen Unterricht zu bestehen scheint, deuten sich weitere sprachliche und kognitive Vorteile zugunsten der bilingual unterrichteten an (Adesope et al., 2010; vgl. Abschnitt 1.5.1). Aufgrund dieser spezifisch im Kontext des bilingualen Sachfachunterrichts zu erwartenden Zusammenhänge wird in der vorliegenden Arbeit, unter Bezugnahme auf Helmkes Angebots-Nutzungs-Modell, der Schülerertrag in Form fachlicher Kompetenzen berücksichtigt. Die Konkretisierung erfolgt domänenspezifisch, also separat für die beiden im bilingualen Unterricht relevanten Fachbereiche der Fremdsprache (Englisch) und des Sachfachs (Geschichte). Bezüglich der Englischkompetenzen wird unter Bezugnahme auf die Ansätze der *unitary* und *divisible competence hypothesis* (Oller & Hinofotis, 1980) die englische Sprachfähigkeit sowohl allgemein, im Sinne einer tiefer liegenden latenten Sprachkompetenz (C-Test), als auch spezifisch, bezogen auf das Hörverstehen als Teilaspekt einer mehrdimensionalen Sprachkompetenz, berücksichtigt. Somit soll sowohl die allgemeine Sprachfähigkeit als auch die tatsächliche Sprachverwendung der Lerner betrachtet werden (vgl. Abschnitt 1.5.2). Bezüglich der Domäne Geschichte wird, in Anlehnung an die Sachkompetenz des

Kompetenzmodells der FUER-Gruppe (Schreiber et al., 2006), die Einsicht der Schülerinnen und Schüler in das Prinzip der Konstruktivität von Geschichte berücksichtigt. Weiterhin wird, unabhängig von den Kompetenzmodellen, das geschichtliche Sachwissen der Lerner in den Blick genommen. In beiden Fällen erscheint es möglich, dass sich aufgrund der bilingualen Unterrichtssituation Vorteile zugunsten der zweisprachig unterrichteten Lerner zeigen. Abgesehen von den bereits weiter oben genannten theoretischen Annahmen, könnten solche Vorteile bezüglich des Erkennens des Konstruktionscharakters von Geschichte vor allem auch durch fremdsprachenbedingte Verfremdungseffekte (Hollm & Pirner, 2008: 117) entstehen, die Lerner leichter dazu in die Lage versetzen, sich den Unterschieden zwischen Quelle und Darstellung und den damit zusammenhängenden Implikationen bewusst zu werden. Weiterhin ermöglicht das Erlernen von Geschichtsinhalten in beiden Sprachen, insbesondere von Fachvokabular und deren zugrunde liegenden Konzepte, eine möglicherweise tiefere Durchdringung des Sachwissens (vgl. Abschnitt 1.5.2).

Die Berücksichtigung von Schülerkompetenzen, wie in den vorherigen Abschnitten dargelegt, stellt den zentralen Aspekt für Analysen zur leistungsbezogenen Wirksamkeit bilingualen Unterrichts dar. Denn dadurch können Aussagen darüber getroffen werden, wie sich die Fähigkeiten von bilingual unterrichteten Schülergruppen über einen bestimmten Zeitraum hinweg entwickeln und ob diesbezüglich Unterschiede im Vergleich zu einsprachig unterrichteten Klassen bestehen. Somit können mögliche Vor- oder Nachteile in der Kompetenzentwicklung bilingual unterrichteter Klassen identifiziert werden. Abschließend soll jedoch darauf hingewiesen werden, dass die Wirksamkeit jeglichen Unterrichts immer in Abhängigkeit davon zu bewerten ist, was zuvor als zu erwartender Ertrag definiert wurde. Das Konzept der Multikriterialität weist nämlich darauf hin, dass die für die vorliegende Arbeit gewählten Schülerkompetenzen nur ein mögliches Kriterium darstellen, anhand dessen die Wirkung von Unterricht beurteilt werden kann (Helmke & Weinert, 1997: 149). Demzufolge könnte bilingualer Geschichtsunterricht auch bezüglich weiterer Faktoren auf seine Wirksamkeit analysiert werden – zum Beispiel daraufhin, wie gut klasseninterne Leistungsunterschiede durch diesen Unterricht reduziert werden können, ob Schülerinnen und Schüler zu Transferleistungen angeregt werden, inwieweit deren Lernstrategien und Arbeitstechniken weiterentwickelt werden oder auf welche Weise bilingualer Unterricht affektive, emotionale, motivationale und soziale Lernziele beeinflusst.

1.6 Bisherige Forschung zur Wirksamkeit bilingualen Unterrichts

1.6.1 Leistungsentwicklung von bilingual unterrichteten Lernern

Studien zur Wirksamkeit von bilingualem Unterricht beschäftigen sich in der Regel mit der Frage, wie sich diese Unterrichtsform auf die Leistungen der Schülerinnen und Schüler auswirkt und wie sich diese im Vergleich zu einsprachig unterrichteten Lernern entwickeln. Hierbei ist es wichtig,

nicht nur den aktuellen Kompetenzstand, sondern auch das Vorwissen der Schülerinnen und Schüler zu erfassen, um auszuschließen, dass ein möglicherweise unterschiedliches Abschneiden der Lernergruppen lediglich an verschiedenen hohen Vorkenntnisständen liegen könnte. Aus diesem Grund können die meisten querschnittlich angelegten Studien keine Aussagen zur Wirksamkeit bilingualen Unterrichts treffen, da ihnen aufgrund ihres einzelnen Messzeitpunkts Einsichten in vorherige Leistungen der Schülerinnen und Schüler verwehrt bleiben. In diesem Sinne liefern Querschnittsstudien zwar beispielsweise wichtige Hinweise darauf, dass bilingualer Unterricht förderlich für die Kompetenzen (Zydati, 2007) und die Motivation (Abendroth-Timmer, 2007) der Schülerinnen und Schüler in der Fremdsprache zu sein scheint sowie zu keinen Nachteilen (Bonnet, 2004) bzw. sogar zu Vorteilen (Heine, 2010; Lamsfuß-Schenk, 2008) im Sachfach führt. Diese Befunde müssen jedoch durch Längsschnittstudien bestätigt werden, die zunächst das Vorwissen der Lerner erfassen und dann mithilfe einer zweiten Messung zu einem späteren Zeitpunkt Aussagen über tatsächliche Veränderungen in den Leistungen treffen können.

Weiterhin können sich mögliche Unterschiede in den Leistungen zwischen bilingual und einsprachig unterrichteten Schülerinnen und Schülern nicht nur aus unterschiedlich großem Vorwissen ergeben, sondern auch aufgrund anderer Faktoren, welche die Schülerleistung beeinflussen, jedoch nicht in direkter Verbindung mit dem bilingualen Unterrichtsangebot stehen. Diese wären, erstens, schüler- bzw. klassenbezogene Lernvoraussetzungen (z. B. sozioökonomischer Status, Intelligenz und Motivation; vgl. Abschnitt 1.2) sowie, zweitens, weitere Aspekte bezüglich derer Unterschiede zwischen bilingual und einsprachig unterrichteten Klassen zu erwarten sind, wie beispielsweise hinsichtlich der Lehrperson (vgl. Abschnitt 1.3), der Unterrichtsqualität (vgl. Abschnitt 1.4.3) und den Leistungsemotionen als Ausdruck der individuellen Wahrnehmung des Unterrichtsangebots durch die Schülergruppen (vgl. Abschnitt 1.4.4). Mögliche Unterschiede in diesen Aspekten sollten in Analysen statistisch konstant gehalten werden, um somit den eigentlichen Effekt des bilingualen Unterrichts auf die Schülerleistungen bestmöglich isolieren zu können (vgl. Helmke & Klieme, 2008: 307).

Um dieses Ziel möglichst idealtypisch umzusetzen, sollten solche Untersuchungen zur Wirksamkeit bilingualen Unterrichts an bilingualen Zug-Schulen (vgl. Abschnitt 1.2.1) die Leistungsentwicklung und die damit in Zusammenhang stehende statistische Kontrolle verschiedenster Aspekte in Bezug auf ganze Klassen durchführen. Ein solches Vorgehen erkennt den bilingualen Unterricht als ein Spezifikum an, das sich auf eine ganze Klasse bezieht und somit ein klassenspezifisches (und nicht schülerspezifisches) Unterscheidungsmerkmal darstellt. Aus diesem Grund entspricht dieses Vorgehen dem forschungsmethodisch gesehen wünschenswerteren Ansatz im Vergleich zu bisherigen Analysen, die sich auf Unterschiede in den Leistungen und Lernvoraussetzungen zwischen bilingual und einsprachig unterrichteten Schülerinnen und Schülern bezogen. Denn diese Untersuchungen nahmen den bilingualen Unterricht nicht als genuin klassenbezogenes Spezifikum in den Blick, sondern lediglich als eine Unterrichtsform, die von

manchen Lernern besucht, von anderen Schülerinnen und Schülern wiederum nicht besucht wird. Folglich wurde die an Schulen typischerweise vorherrschende Klassenstruktur – nach der es keine „einzelnen“ Lerner gibt, da diese immer Klassen zugeordnet sind – in diesen Studien nicht berücksichtigt. Dies kann zu verzerrten Ergebnissen führen, da die fehlende Berücksichtigung der geklumpte Stichprobe (d. h. Lerner geschachtelt in Klassen) mit großer Wahrscheinlichkeit eine Überschätzung der Effekte bewirkt (Raudenbush & Bryk, 2002).

Weiterhin kann die leistungsbezogene Wirksamkeit bilingualen Unterrichts nicht nur auf Basis der weiter oben beschriebenen Vergleiche mit einsprachigem Sachfachunterricht bestimmt werden, welche Auskunft darüber geben, ob sich möglicherweise spezifische Vor- oder Nachteile für bilingual unterrichtete Klassen in der Leistungsentwicklung ergeben. Dies kann auch mithilfe der Berücksichtigung von spezifischen methodisch-didaktischen Aspekten erfolgen, die dieser Unterrichtsform inhärent sind und anhand deren sich das bilinguale Unterrichtsangebot von anderen Unterrichtsformen unterscheidet, wie beispielsweise der Einsatz von Mutter- und Fremdsprache (vgl. Abschnitte 1.4.1 und 1.4.2). Solche Analysen zu spezifischen Wirkungsweisen beziehen sich lediglich auf das bilinguale Unterrichtssetting und untersuchen, wie solche speziellen methodisch-didaktischen Maßnahmen dieser Unterrichtsform die Ausbildung der Kompetenzen in den bilingual unterrichteten Klassen bedingen könnte.

In Anknüpfung an diese Überlegungen soll im Folgenden die bisherige Forschung zur Leistungsentwicklung von Schülerinnen und Schülern im bilingualen Unterricht vorgestellt werden. Hierfür werden zunächst Längsschnittstudien zitiert, die Vorkenntnisse und zum Teil auch individuelle Lernvoraussetzungen der Schülerinnen und Schüler berücksichtigten (Abschnitt 1.6.2). Daran schließen sich Forschungsarbeiten zu weiteren Aspekten an, die die Schülerleistung beeinflussen und im Kontext des bilingualen Unterrichts von besonderer Relevanz zu sein scheinen (Lehrperson, Unterrichtsqualität, Leistungsemotionen; Abschnitt 1.6.3). Abschließend folgen Studien zu spezifischen methodisch-didaktischen Unterrichtsmerkmalen, welche diese Leistungen möglicherweise bedingen (Abschnitt 1.6.4). Die folgenden Abschnitte sollen hierbei eine systematische und überblicksartige Darstellung von für den vorliegenden Kontext relevanten Forschungsarbeiten verkörpern. Sie zielt darauf ab, die Haupteckdaten der Studien wiederzugeben und auf Besonderheiten, Gemeinsamkeiten und Unterschiede in den Befunden hinzuweisen. Zu diesem Zweck wird auf eine detaillierte Darstellung der jeweils zugrunde liegenden Designs und der genauen Ergebniswerte verzichtet. Ausführlichere Informationen finden sich in den einzelnen Teilstudien (Kapitel 3 bis 6).

1.6.2 Studien zur Leistungsentwicklung im bilingualen Unterricht

Bisher liegen bereits mehrere längsschnittliche Studien zu Englischkompetenzen im bilingualen Unterricht vor. Die Arbeit von Bos, Bonsen und Gröhlich (2009) im Rahmen der KESS-Studie nimmt hierbei eine besondere Stellung ein. Sie erhob die Englischkompetenzen von mehr als 2000

Schülerinnen und Schülern bereits zum Ende der Grundschulzeit, bevor diese in einer weiterführenden Schule entweder an einem bilingualen Zug oder an einer anderen Spezialisierung der Schulen teilnahmen (vgl. Abschnitt 1.2.1). Bos et al. (2009) stellten fest, dass die zukünftigen Lerner der bilingualen Züge bereits am Ende der 4. Klasse über bessere Englischkompetenzen verfügten als diejenigen Schülerinnen und Schüler, die sich für ein anderes Profil entschieden hatten. Weiterhin konnten sie zeigen, dass sich dieser Leistungsvorsprung der Lerner aus bilingualen Zügen bis zum Ende der 6. Klasse weiter vergrößerte, was vermutlich auf den verstärkten Englischunterricht in den Klassenstufen 5 und 6 zurückgeführt werden kann (vgl. Abschnitt 1.2.1). Zusammenfassend zeigten die Schülerinnen und Schüler der bilingualen Züge also deutlich bessere Englischleistungen als diejenigen Lerner, die sich nicht für einen bilingualen Zug entschieden hatten. Hierbei ist zu unterstreichen, dass sich dieser Leistungsvorsprung lediglich auf Daten bezog, die zum Ende der Grundschulzeit und im Anschluss an den zusätzlichen Englischunterricht in den Klassen 5 und 6 erhoben wurden. Die Effekte des bilingualen Unterrichts an sich, der erst ab Klassenstufe 7 erteilt wurde, waren darin noch gar nicht berücksichtigt. Folglich scheint es in Anlehnung an Krashens (1985) Input-Hypothese sowie weitere Zweitspracherwerbtheorien (vgl. Abschnitt 1.5.1) sehr wahrscheinlich, dass sich diese bereits bestehenden Vorteile der Schülerinnen und Schüler in ihren Englischkompetenzen noch weiter vergrößerten im weiteren Verlauf ihrer Teilnahme am bilingualen Zug.

Zu ähnlichen Ergebnissen kam auch Fehling (2008), wobei angemerkt werden muss, dass sie mögliche Effekte des verstärkten Englischunterrichts der Klassen 5 und 6 nicht erfassen konnte. Sie untersuchte die *language awareness* von 139 bilingual und einsprachig unterrichteten Schülerinnen und Schülern im Verlauf der 7. und 8. Klassenstufe, nachdem der bilinguale Unterricht einsetzte. Sie kam zu dem Ergebnis, dass die bilingual unterrichteten Lerner bereits zu Beginn des bilingualen Unterrichts, Anfang der 7. Klasse, eine bessere *language awareness* aufwiesen und sie diese Fähigkeit im Verlauf der beiden Schuljahre noch stärker ausbauen konnten, als die einsprachig unterrichteten Schülerinnen und Schüler. Diese Befunde werden bestätigt von Alonso, Grisaleña und Campo (2008) sowie Bredenbröcker (2000), die ebenfalls günstigere Leistungsentwicklungen der bilingual unterrichteten Lerner in Englisch im Vergleich zu den einsprachig unterrichteten Schülerinnen und Schülern feststellten. Bei Alonso et al. (2008) bezogen sich die Ergebnisse auf 229 Lerner, die in der Sekundarstufe im Abstand von 1,5 Jahren zwei Mal getestet wurden. Bredenbröckers (2000) Analysen basierten auf den Daten von 195 bilingual und einsprachig unterrichteten Schülerinnen und Schülern, die – wie bei Fehling (2008) – drei Mal im Verlauf der 7. und 8. Klasse an Erhebungen teilnahmen.

Weiterhin ist eine Studie bekannt, die die Leistungsentwicklung im bilingual unterrichteten Sachfach in den Blick nahm. Jäppinen (2005) untersuchte die mathematischen und naturwissenschaftlichen Fähigkeiten von 669 Grund- und Sekundarstufenschülerinnen und -schülern zu vier Zeitpunkten in halbjährlichen Intervallen. Im Gegensatz zu den Fremdsprachenkompetenzen

(s. oben) zeigten sich für die Sachfachkompetenzen zu Beginn der Erhebungen keine Anfangsunterschiede zwischen den Gruppen. Weiterhin verlief die Leistungsentwicklung der bilingual und einsprachig unterrichteten Lerner über die zwei Jahre hinweg vergleichbar. Für bilingual Unterrichtete ergaben sich also bezüglich dieser Domäne keine Leistungsvorteile. Angesichts der wenigen längsschnittlichen Studien zur Kompetenzentwicklung von Schülerinnen und Schülern im bilingual unterrichteten Sachfach (Bonnet, 2012: 67; Bernhardt, 2011: 215; Coyle et al., 2010: 138; Sylvén, 2010: 224), können diese Befunde bisher kaum durch weitere Analysen bestätigt oder widerlegt werden.

Die Gemeinsamkeit aller bisher zitierten Studien liegt jedoch darin, dass die individuellen Lernvoraussetzungen der Schülerinnen und Schüler nicht berücksichtigt wurden, welche jedoch die Schülerleistung beeinflussen. So erhob Fehling (2008) zwar die verbalen kognitiven Grundfähigkeiten der Lerner und stellte fest, dass diese bei den bilingual unterrichteten Schülerinnen und Schülern besser ausfielen. Sie setzte diese aber nicht in Bezug zur Leistungsentwicklung in Englisch. Somit konnte Fehling (2008) nicht zeigen, ob die fremdsprachliche Leistungsentwicklung der bilingual Unterrichteten unter Berücksichtigung ihrer Vorteile in den verbalen Fähigkeiten möglicherweise weniger positiv verlaufen wäre. Dasselbe gilt für die Studien von Alonso et al. (2008) und Bredenbröker (2000), die neben Fremdsprachenkompetenzen auch die Einstellung der Lerner zu Englisch erhoben hatten. Obwohl diese bei den bilingual Unterrichteten deutlich positiver ausfiel, wurde die Einstellung zur Fremdsprache bei der Analyse der Leistungsentwicklung nicht miteinbezogen. Somit hätten die Unterschiede in den Leistungen zwischen bilingual und einsprachig unterrichteten Schülerinnen und Schülern, die sich in den Studien zeigten, besser erklärt werden können. Im Folgenden werden nun Forschungsarbeiten vorgestellt, die diese Lernvoraussetzungen der Schülerinnen und Schüler berücksichtigt haben.

Köller et al. (2012) erfassten die Englischkompetenzen von knapp 1000 bilingual und einsprachig unterrichteten Neuntklässlern unter statistischer Kontrolle von Geschlecht, Migrationshintergrund, Deutschfähigkeiten und sozioökonomischem Status. Die Berücksichtigung dieser Schülermerkmale bewirkte jedoch lediglich eine geringe Reduktion des Leistungsvorteils der bilingual unterrichteten Lerner in den Englishtests, da sich beide Lernergruppen, insbesondere hinsichtlich des sozioökonomischen Status, kaum voneinander unterschieden. D. h., die bilingual Unterrichteten zeigten trotz Kontrolle einiger Lernvoraussetzungen bessere Leistungen als die einsprachig unterrichteten Lerner. Allerdings kann dieser Befund nur unter Vorbehalt betrachtet werden. Denn abgesehen von der wichtigen Berücksichtigung individueller Schülermerkmale ist einschränkend anzumerken, dass Köller et al. (2012) lediglich zu einem Messzeitpunkt Daten erhoben haben und somit die fremdsprachlichen Vorkenntnisse der Lerner nicht berücksichtigen konnten, die eine zentrale Rolle bei Wirksamkeitsanalysen von Unterricht spielen (vgl. Abschnitt 1.6.1).

Badertscher und Bieri (2009) hingegen nutzten ein längsschnittliches Design, um die Kompetenzentwicklung bilingual unterrichteter Schülerinnen und Schüler aus sechs Sekundarschulklassen im Sachfach Geschichte zu analysieren. In ihrer Studie bestand keine Notwendigkeit zur Berücksichtigung individueller Lernvoraussetzungen. Denn Badertscher und Bieri (2009) setzten ein *within-subject* Design ein, wonach dieselben Schülerinnen und Schüler zwei Mal im Abstand von 1,5 Jahren sowohl an einer bilingual als auch an einer einsprachig gehaltenen Unterrichtseinheit teilnahmen und jeweils vor und nach den Modulen einen Wissenstest absolvierten. Dieses Design wies den Vorteil auf, dass keine Notwendigkeit zur Berücksichtigung möglicher Unterschiede in Schülermerkmalen bestand, da die Lerner mit sich selbst verglichen wurden. Allerdings sind die Befunde nur sehr begrenzt aussagefähig, da das Unterrichtssetting (bi- oder monolingual) zu hundert Prozent mit dem jeweiligen Modulthema konfundiert war. Dennoch verwiesen die Ergebnisse von Badertscher und Bieri (2009) – wie die von Jäppinen (2005) – auf vergleichbare Leistungsentwicklungen der Schülergruppen. Sie stellten fest, dass sich die Qualität des in den Modulen hinzugelernten Wissens nicht unterschied – egal, ob die Unterrichtseinheiten auf Deutsch oder bilingual unterrichtet wurden.

Bei Nold et al. (2008) hingegen bestanden die obigen forschungsmethodisch bedingten Einschränkungen nicht. Die Forschergruppe analysierte die Entwicklung der Englischkompetenzen mit einer Teilstichprobe von DESI, die aus knapp 2000 Schülerinnen und Schülern der 9. Klassenstufe bestand. Sie ordneten die bilingual und einsprachig unterrichteten Lerner einander so zu, dass diese die gleichen Merkmale bzw. ähnliche Ausprägungen in den folgenden Bereichen zeigten: Geschlecht, Deutschkompetenzen, kognitive Grundfähigkeiten, sozioökonomischer Status. Nold et al. (2008) konnten zeigen, dass die bilingual unterrichteten Lerner trotz dieser vergleichbaren Ausgangsbedingungen über ein höheres Englischniveau verfügten als die einsprachig unterrichteten Schülerinnen und Schüler und darüber hinaus im Hörverstehen deutlich größere Leistungszuwächse zeigten. Diese größeren Zuwächse zeigten sich jedoch nicht bezüglich der allgemeinen englischen Sprachfähigkeit der Lerner (DESI-Konsortium, 2006).

Rumlich (2012) kam zu ähnlichen Ergebnissen. In der DENOCS-Studie erhob er den Leistungszuwachs in Englisch von 626 bilingual und einsprachig unterrichteten Schülerinnen und Schülern im Verlauf der 6. bis 8. Klassenstufe unter statistischer Kontrolle von Geschlecht, verbalen kognitiven Grundfähigkeiten, Interesse und akademischem Selbstkonzept in Englisch. Die einsprachig Lernenden setzten sich aus zwei Gruppen zusammen. Die Schülerinnen und Schüler der einen Vergleichsgruppe stammten aus den Parallelklassen der bilingual unterrichteten Lerner. Diejenigen der anderen Vergleichsgruppe besuchten Schulen ohne bilinguales Unterrichtsangebot. Rumlich (2013) stellte fest, dass die bilingual unterrichteten Schülerinnen und Schüler bereits zu Beginn der 6. Klasse über bessere Englischkompetenzen sowie mehr Interesse und ein höheres Selbstkonzept (Rumlich, 2014a) verfügten als die beiden Vergleichsgruppen. Weiterhin entwickelten sich ihre Leistungen bis zum Ende der 8. Klasse günstiger als die der Lerner aus den Parallelklassen.

Bezüglich der zweiten Vergleichsgruppe jedoch, Schülerinnen und Schüler aus Schulen ohne bilinguales Unterrichtsangebot, zeigten sich vergleichbare Kompetenzentwicklungen mit den bilingual unterrichteten Lernern (Rumlich, 2014b). Dieser Befund verweist auf Nachteile für diejenigen Schülerinnen und Schüler, die sich an einer Schule mit bilingualem Zug gegen diesen, und für ein anderes Profil, entschieden haben. Da Rumlichs Studie jedoch, zusammen mit den querschnittlichen Untersuchungen von Daniel (2001) zu Lexik im bilingualen Unterricht, einer der wenigen Vorarbeiten zum Thema verschiedener einsprachig unterrichteter Vergleichsgruppen darstellt, bedarf es weiterer Forschung, um diesen Aspekt vertieft zu analysieren.

Anknüpfend an Rumlichs (2012) Studie und unter Verweis auf Fehlings (2008) Arbeit ist weiterhin hervorzuheben, dass die verbalen Fähigkeiten von Schülerinnen und Schülern als wichtiger Bestandteil der Lernvoraussetzungen in Erhebungen zum bilingualen Unterricht zwar bereits Berücksichtigung fanden. Es zeigte sich, dass bilingual Unterrichtete diesbezüglich bessere Leistungen erbrachten als einsprachig unterrichtete Lerner. Dies könnte möglicherweise an deren sprachbezogener Begabung und verstärktem Interesse an Sprachen liegen (vgl. Abschnitt 1.2.2). Was jedoch bisher bei vergleichenden Analysen zwischen bilingual und einsprachig unterrichteten Lernern kaum untersucht wurde, sind andere kognitive Fähigkeitsbereiche, wie beispielsweise die figuralen kognitiven Grundfähigkeiten. Weiterhin sind – wie bereits weiter oben erwähnt – zurzeit keine Studien bekannt, die die Leistungsentwicklung bilingual unterrichteter Schülerinnen und Schüler im Sachfach unter Berücksichtigung individueller Lernvoraussetzungen in den Blick genommen haben. Dies stellt ein bedeutsames Forschungsdesiderat dar, angesichts dessen, dass die Debatte um den „Mehrwert“ bilingualen Unterrichts für die Sachfächer im letzten Jahrzehnt ausgiebig und noch nicht abschließend geführt wurde (Hollm, 2013: 185; Breidbach & Viebrock, 2012: 8; Theis, 2010; Barricelli & Schmieder, 2009; Hasberg, 2004). Schließlich sind keine Studien bekannt, die bei Analysen zur Wirksamkeit dieser Unterrichtsform die Leistungsentwicklungen von ganzen bilingual und einsprachig unterrichteten Klassen miteinander verglichen haben. Dies trifft ebenfalls auf die Berücksichtigung möglicher Unterschiede in deren durchschnittlichen Lernvoraussetzungen zu, die in entsprechenden Analysen auf Klassenebene statistisch kontrolliert werden müssten (vgl. Abschnitt 1.6.1). Dies stellt ein weiteres Forschungsdesiderat dar, das es zu schließen gilt.

1.6.3 Studien zur Lehrperson, Unterrichtsqualität und Schüleremotionen im Kontext des bilingualen Unterrichts

In den Abschnitten 1.3 und 1.4 wurden drei Faktoren (Lehrperson, Unterrichtsqualität und Leistungsemotionen) identifiziert, bezüglich derer Unterschiede zwischen bilingualem und einsprachigem Unterricht angenommen werden. Da diese – wie auch die individuellen und klassenbezogenen Lernvoraussetzungen (vgl. Abschnitt 1.6.2) – die fachlichen Leistungen der Schülerinnen und Schüler beeinflussen können (Helmke, 2012), ist es wichtig, sie im Rahmen von

Wirksamkeitsanalysen zu berücksichtigen, um den tatsächlichen Effekt des bilingualen Unterrichts auf die Schülerkompetenzen bestmöglich isolieren zu können (vgl. Abschnitt 1.6.1). Im Folgenden werden deshalb Forschungsarbeiten vorgestellt, die sich mit der Unterrichtsqualität, der Lehrperson und den Leistungsemotionen der Lerner im Kontext des bilingualen Unterrichts beschäftigt haben.

Unterrichtsqualität. Meyer (2003) untersuchte mithilfe von Fragebögen und Interviews die Einschätzung, Wahrnehmung und Bewertung des bilingualen Geographieunterrichts von rund 500 bilingual und einsprachig unterrichteten Schülerinnen und Schülern der Klassenstufen 8 bis 10. Obwohl Meyer (2003) nicht explizit angab, Unterrichtsqualität zu erfassen, deuteten ihre Ergebnisse an, dass diese im bilingualen Unterricht positiver ausfiel. Demnach schienen die zweisprachig unterrichteten Lerner mehr Unterstützung im Unterricht durch die Lehrkraft zu erfahren, in mehr schüleraktivierenden Tätigkeiten eingebunden gewesen zu sein sowie von interessanteren Lehrmethoden profitiert zu haben als die einsprachig unterrichteten Schülerinnen und Schüler. Über diese Untersuchung von Meyer (2003) hinaus sind jedoch keine weiteren Studien bekannt, die sich im Rahmen des bilingualen Unterrichts mit dem Thema der Unterrichtsqualität beschäftigt haben.

Lehrperson. Zur bilingual unterrichtenden Lehrperson wurden im letzten Jahrzehnt mehrere Forschungsarbeiten durchgeführt. Diese nahmen oft (schul-)biographische Aspekte der Lehrkräfte in den Blick, wie beispielsweise Informationen zum eigenen ein- oder mehrsprachigen Hintergrund, zu eigenen Auslandsaufenthalten im Zielsprachenland, zur Aus- und Weiterbildung als bilinguale Lehrkraft oder die Anzahl bisher bilingual unterrichteter Schuljahre (Kollenrott, 2008; Müller-Bittner, 2008; Gierlinger et al., 2007; Viebrock, 2007; Zydariß, 2007; Helbig, 2001). Zu denjenigen Aspekten jedoch, die in Abschnitt 1.3 als relevant für den vorliegenden Kontext identifiziert wurden (Motivation, Professionswissen sowie allgemeine schul- und unterrichtsbezogene Erwartungen, Ziele und Einstellungen bilingual unterrichtender Lehrkräfte), wurden bisher nur wenige Analysen durchgeführt. In Meyers (2003) Studie wurde, abgesehen vom Aspekt der Unterrichtsqualität, auch Motivation erhoben. Hierbei wurden die bilingual und einsprachig unterrichteten Schülerinnen und Schüler gebeten, ihre Lehrkraft diesbezüglich einzuschätzen. Die Ergebnisse verwiesen darauf, dass bilingual unterrichtende Lehrpersonen im Vergleich zu ihren einsprachig unterrichtenden Kolleginnen und Kollegen besonders motiviert und motivierend gewesen zu sein schienen. Ähnliche Befunde wies D'Angelo (2013) auf. Sie analysierte Fragebögen von 215 bilingual unterrichtenden Sekundarschullehrkräften und kam zu dem Ergebnis, dass sich diese in bilingual gehaltenen Unterrichtsstunden als motivierter erlebten im Vergleich zu einsprachig gehaltenen Schulstunden und dort immer „100% ihrer Möglichkeiten“ genutzt hätten (D'Angelo, 2013: 112). Abgesehen von diesen Befunden zur Motivation bilingual unterrichtender Lehrkräfte sind keine weiteren Studien bekannt, die sich mit deren Professionswissen oder allgemeinen schul- und unterrichtsbezogenen Erwartungen, Zielen und Einstellungen beschäftigten; mit denjenigen Aspekten also, bezüglich derer Unterschiede zu einsprachig unterrichtenden Lehrkräften erwartet werden (vgl. Abschnitt 1.3.2). In diesem Zusammenhang soll unterstrichen werden, dass auch die bereits bestehenden Studien zur

bilingual unterrichtenden Lehrperson, wie etwa die von Meyer (2003) oder D'Angelo (2013), lediglich solche Aussagen treffen konnten, die auf Schülerurteilen oder Angaben der bilingual unterrichtenden Lehrerinnen und Lehrer basierten. Was bisher jedoch noch nicht durchgeführt wurde, sind vergleichende Analysen, denen Informationen sowohl zu bilingual als auch einsprachig unterrichtenden Lehrkräften zugrunde liegen. Dadurch könnten Unterschiede zwischen diesen beiden Lehrergruppen systematisch analysiert und darauffolgend in Zusammenhang mit Schülerleistungen gebracht werden.

Emotionen. Im bilingualen Unterricht wurden die Leistungsemotionen der Schülerinnen und Schüler als Ausdruck ihrer individuellen Wahrnehmung des Unterrichtsangebots bisher kaum beforscht. Es sind lediglich zwei Studien bekannt, die Angst, Ärger und Freude (Piesche, Jonkmann & Keßler, 2014) bzw. Freude und Anspannung (Hartmannsgruber, 2014) der Lerner im bilingualen Unterricht in den Blick nahmen und mit den Emotionen von einsprachig unterrichteten Schülerinnen und Schülern verglichen. Hartmannsgruber (2014) stellte keine statistisch signifikant unterschiedlichen Entwicklungen der Emotionen zwischen den beiden Schülergruppen fest. Die Ergebnisse von Piesche et al. (2014) verwiesen darauf, dass die bilingual Unterrichteten weniger positive und mehr negative Emotionen im Zusammenhang mit dem bilingualen Unterrichtsangebot erlebten als die einsprachig unterrichteten Schülerinnen und Schüler. Diese Befunde müssen jedoch mit Vorbehalt rezipiert werden, da ihnen andersartige Stichproben zugrunde liegen. Während in der vorliegenden Arbeit Schülerinnen und Schüler aus fest etablierten bilingualen Zügen im Mittelpunkt stehen (vgl. Abschnitt 1.2.1), fanden bei Hartmannsgruber (2014) und Piesche et al. (2014) die Erhebungen mithilfe von Klassen statt, die per Zufall für die Teilnahme an einem bilingual oder einsprachig unterrichteten Modul ausgewählt wurden. Diese Schulen besaßen keinen bilingualen Zug und kein(e) Schüler(in) hatte zuvor Erfahrungen mit bilinguaem Unterricht gesammelt. Aufgrund dieser variierenden Ausgangsbedingungen können die Ergebnisse der beiden Studien nicht auf den Kontext der vorliegenden Arbeit angewandt werden. Diese geht von einer positiv selektierten Schülerschaft aus (vgl. Abschnitt 1.2) und erwartet aufgrund dessen positivere Emotionen der bilingual unterrichteten Lerner (vgl. Abschnitt 1.4.4). Es sind bisher keine Studien bekannt, die sich innerhalb dieser Rahmenbedingungen mit Leistungsemotionen beschäftigt haben.

Abschließend soll darauf hingewiesen werden, dass alle in diesem Abschnitt zitierten Forschungsarbeiten zwar Aspekte zu Unterrichtsqualität, Lehrperson und Emotionen im Kontext des bilingualen Unterrichts untersucht und deskriptiv darüber berichtet haben. Sie wurden jedoch nicht in Beziehung gesetzt mit den Schülerleistungen und daraufhin untersucht, wie sich diese auf den fachlichen Unterrichtsertrag auswirken könnten. Die Berücksichtigung solcher Faktoren, die zwar die Schülerkompetenzen mitbeeinflussen (vgl. Helmke, 2012), jedoch nicht in direktem Zusammenhang mit dem bilingualen Unterrichtsangebot stehen, stellt ein wichtiges Vorgehen bei Forschungsarbeiten zur Wirksamkeit dieser Unterrichtsform dar und sollte in entsprechenden Studien Anwendung finden (vgl. Abschnitt 1.6.1).

1.6.4 Studien zu methodisch-didaktischen Aspekten des bilingualen Unterrichts

In jedem Unterricht werden fortwährend methodisch-didaktische Maßnahmen seitens der Lehrkräfte umgesetzt, die bei Schülerinnen und Schülern zu einer gewünschten Veränderung führen sollen. Diese beziehen sich oft auf eine Verbesserung der Leistungen. Für den vorliegenden Kontext des bilingualen Unterrichts sind Studien von Interesse, die solche spezifischen, mit dieser Unterrichtsform in Verbindung stehende Maßnahmen beschrieben sowie deren mögliche Auswirkungen auf Schülerleistung untersucht haben. Für solche Analysen, die nur das bilinguale Unterrichtssetting betreffen, bedarf es demnach keiner vergleichenden Befunde aus einsprachigem Sachfachunterricht.

Bisher wurden bereits mehrere Studien durchgeführt, die verschiedene spezifische methodisch-didaktische Aspekte des bilingualen Unterrichts in den Blick genommen haben. Viele Untersuchungen beschäftigten sich dabei mit Sozialformen sowie dem Einsatz von Unterrichtsmaterial und Schulbüchern im bilingualen Unterricht (Kollenrott, 2008; Müller-Bittner, 2008; Gierlinger et al., 2007; Viebrock, 2007; Müller-Schneck, 2006; Meyer, 2003; Helbig, 2001). Dies war wohl nicht zuletzt deshalb der Fall, da diesbezüglich von einigen betreffenden Ministerien sowie in der Fachdidaktik (z. B. KM, 2006; 9; Schönfeldt, 2005: 148f.) Empfehlungen ausgesprochen wurden, wonach im bilingualen Unterricht verstärkt Gruppen- und Partnerarbeit sowie mehr Visualisierungsformen eingesetzt werden sollten. Als Resümee der eben zitierten Studien kann zusammengefasst werden, dass den Angaben der befragten Lehrkräfte zufolge im bilingualen Unterricht tatsächlich häufiger Gruppen- und Partnerarbeitsphasen zum Einsatz kamen und größerer Wert auf Visualisierung gelegt wurde. Weiterhin schienen die bilingual unterrichtenden Lehrkräfte im Allgemeinen sehr unzufrieden mit der mangelhaften Materiallage und einige gaben an, kein Schulbuch im bilingualen Unterricht zu verwenden. Abgesehen von diesem Fokus auf Sozialformen und Materialeinsatz befindet sich jedoch auch die Verwendung von Mutter- und Fremdsprache im bilingualen Unterricht in der aktuellen fachdidaktischen Diskussion (Diehr, 2012: 29; vgl. Abschnitt 1.4.2). Dieser Aspekt soll in Anlehnung an die Überlegungen des Abschnitts 1.4.1, die zeigten, dass dies das Spezifikum dieser Unterrichtsform darstellt, im Folgenden forschungsbezogen vertieft werden.

Kollenrott (2008), Gierlinger et al. (2007), Zydati (2007), Müller-Schneck (2006) und Bredenbröker (2000) thematisierten das grundsätzliche Verhältnis zwischen fremd- und muttersprachlichem Redeanteil im bilingualen Unterricht. Mithilfe von Lehrerfragebögen und Interviews (Stichprobengrößen zwischen 10 und 134 Personen) konnten die Studien unter großer Übereinstimmung zeigen, dass die überwiegende Mehrheit der Lehrkräfte angab, größtenteils die Fremdsprache zur Kommunikation im Unterricht einzusetzen. Die Muttersprache wurde den Lehrkräften zufolge jedoch nie komplett ausgeschlossen, sondern zumindest manchmal, oft im Sinne eines „Verständnis-Notnagels“, verwandt. Ähnliches traf auch auf den Gebrauch der beiden

Sprachen durch die Lerner zu (Gierlinger et al., 2007). Weitere Arbeiten beschäftigten sich mit der Frage, auf welche Weise das Deutsche von den Lehrkräften eingesetzt wurde und kamen zu dem Ergebnis, dass dies immer zweckgebunden, in bestimmten Situationen zur Erreichung bestimmter Ziele, der Fall war. So gaben die Lehrkräfte beispielsweise an, Deutsch zur Erklärung komplizierter Sachverhalte, zur Absicherung des Sachwissens durch Wiederholungen und Zusammenfassungen sowie bei der Einführung neuer Fachbegriffe zu verwenden (Hüttner & Dalton-Puffer, 2013; Hollm, 2011; Gierlinger et al., 2007; Meyer, 2003; Bredenbröker, 2000). Bezüglich dieses letzten Aspekts liegen jedoch keine detaillierten Informationen darüber vor, auf welche Weise Fachbegriffe zwei- und/oder einsprachig unterrichtet wurden. Dies überrascht angesichts der Tatsache, dass Fachtermini im bilingualen Unterricht eine zentrale Stellung einnehmen (Heimes, 2013: 349) und sowohl in der Mutter- als auch in der Fremdsprache eingeführt werden sollten (Theis, 2010: 52; Otten & Wildhage, 2007: 19; KM, 2006: 9). Als weitere deutsche Sprechanlässe wurden fehlendes fremdsprachliches Material (Müller-Bittner, 2008), Zeitmangel (Viebrock, 2007), Themen im Geschichtsunterricht mit vornehmlichem Bezug zur deutschen Geschichte (Müller-Schneck, 2006) sowie Gruppen- und Erarbeitungsphasen (Helbig, 2001) genannt.

Zusammenfassend liegen nur einzelne und wenig differenzierte Befunde zu diesem entscheidenden Merkmal des bilingualen Unterrichts vor. Darüber hinaus ist es wichtig anzumerken, dass die sprachbezogenen methodisch-didaktischen Aspekte in allen angeführten Studien lediglich deskriptiv vorgestellt, jedoch nicht in Bezug zu den Leistungen der bilingual unterrichteten Klassen gesetzt wurden (Hollm, 2011: 329; vgl. Fazit des vorhergehenden Abschnitts 1.6.3.). Somit konnten bisher noch keine Aussagen darüber getroffen werden, wie sich der Einsatz von Mutter- und Fremdsprache im bilingualen Unterricht auf die Sprach- und Sachfachkompetenzen der Schülerinnen und Schüler auswirken könnte.

1.6.5 Zusammenfassung und Diskussion

Bisher wurden bereits einige längsschnittliche Studien durchgeführt, die die Leistungen von bilingual unterrichteten Schülerinnen und Schülern in den Blick nahmen. Obige Darstellungen konnten zeigen, dass bilingual unterrichtete Schülerinnen und Schüler bereits zu Beginn des bilingualen Unterrichts in der 7. Klasse, wahrscheinlich aber schon zum Ende der Grundschule, also vor Beginn des verstärkten Englischunterrichts in den Klassen 5 und 6, über bessere Englischkompetenzen verfügen als einsprachig unterrichtete Lerner. Abgesehen von diesem Vorteil scheinen die bilingual unterrichteten mit grundsätzlich günstigeren Lernvoraussetzungen in die weiterführenden Schulen einzutreten, da diese höhere verbale kognitive Fähigkeiten sowie mehr Interesse und eine positivere Einstellung zur Fremdsprache aufzuweisen scheinen. Bezüglich des sozioökonomischen Status, figuraler kognitiver Fähigkeiten sowie der Motivation und den Kenntnisständen der Schülerinnen und Schüler im bilingual unterrichteten Sachfach verweisen die wenigen – wenn überhaupt – vorliegenden Studien auf ähnliche Ausgangslagen der beiden Lernergruppen. Was die tatsächliche

Leistungsentwicklung der Schülerinnen und Schüler im Verlauf des bilingualen Unterrichts betrifft, weist die bisherige Forschung darauf hin, dass diese für die bilingual Unterrichteten in Englisch deutlich positiver ausfällt als für die einsprachig unterrichteten Schülerinnen und Schüler, und zwar vor allem im Bereich der rezeptiven Fähigkeiten wie beispielsweise im Hörverstehen (vgl. Abschnitt 1.5.2). Dies scheint auch dann zuzutreffen, wenn Unterschiede in den individuellen Lernvoraussetzungen statistisch kontrolliert, die günstigeren Ausgangslagen der bilingual unterrichteten Lerner also berücksichtigt wurden. Für das Sachfach hingegen konnten die wenigen vorliegenden Studien vergleichbare Kompetenzentwicklungen der beiden Lernergruppen aufzeigen. Diese Untersuchungen bezogen jedoch keine individuellen Schülermerkmale mit ein, was tendenziell in einer etwas schlechteren Leistungsentwicklung seitens der bilingual unterrichteten Lerner resultieren könnte. Schließlich sind keine Studien bekannt, die die Leistungsentwicklung ganzer bilingual und einsprachig unterrichteter Klassen miteinander verglichen und dabei mögliche Unterschiede in den Lernvoraussetzungen und Vorkenntnisständen auf Klassenebene statistisch kontrolliert haben. Somit wurde dieser forschungsmethodisch gesehen präzisere Ansatz, der den bilingualen Unterricht als ein genuin klassenspezifisches Merkmal erfasst, in bisherigen Studien noch nicht umgesetzt.

Für Wirksamkeitsanalysen bietet es sich jedoch an, nicht nur Vorkenntnisse und Lernvoraussetzungen der Schülerinnen und Schüler bzw. Klassen statistisch konstant zu halten. Um die Effekte des bilingualen Unterrichts bestmöglich isolieren zu können, sollten auch weitere, die Schülerleistung beeinflussenden Aspekte berücksichtigt werden, die als besonders relevant für den vorliegenden Kontext identifiziert wurden, wie beispielsweise Unterrichtsqualität, Lehrperson und individuelle Wahrnehmung des Unterrichtsangebots (vgl. Abschnitte 1.3, 1.4.3, 1.4.4). Die wenigen Studien, die hierzu bisher bekannt sind, deuteten an, dass der bilinguale Unterricht eine höhere Qualität aufweisen könnte als der einsprachig gehaltene Sachfachunterricht. Weiterhin scheinen bilingual unterrichtende Lehrkräfte motivierter und motivierender zu sein als ihre einsprachig unterrichtenden Kolleginnen und Kollegen. Zu anderen lehrpersonbezogenen Aspekten, wie beispielsweise Professionswissen oder allgemeine schul- und unterrichtsbezogene Einstellungen, Ziele und Erwartungen, sind keine Untersuchungen bekannt. Bezüglich der individuellen Wahrnehmung des Unterrichtsangebots in Form von Leistungsempfindungen der Schülerinnen und Schüler liegen bisher ebenfalls keine Befunde vor, die auf den vorliegenden Kontext von Lernern in bilingualen Zügen übertragbar wären. Darüber hinaus wurden im Sinne einer umfassenden Wirksamkeitsanalyse des bilingualen Unterrichts bisher keine dieser die Schülerleistung beeinflussenden Aspekte in Bezug gesetzt mit der Kompetenzentwicklung der Klassen.

Weiterhin wurden einige Studien vorgestellt, die sich mit den spezifischen methodisch-didaktischen Aspekten des bilingualen Unterrichts beschäftigten. Dabei sind für die vorliegende Arbeit solche Untersuchungen relevant, die das Spezifikum des bilingualen Unterrichts, den Einsatz der Mutter- und Fremdsprache, in den Blick nahmen. Bisherige Forschungsarbeiten konnten zeigen,

dass die Fremdsprache die überwiegende Kommunikationssprache im bilingualen Unterricht darzustellen scheint, die Muttersprache jedoch auch zweckgebundene Berücksichtigung findet. Dies scheint meist bei Verständnisschwierigkeiten seitens der Schülerinnen und Schüler der Fall zu sein, also zur Erklärung komplizierter Sachverhalte, zur Absicherung des Sachwissens durch Wiederholungen und Zusammenfassungen sowie bei der Einführung neuer Fachbegriffe. Diese Befunde sind nicht nur im Sinne ihrer deskriptiven Beschreibung von Interesse, sondern können darüber hinaus auch interessante Erkenntnisse hinsichtlich der Wirksamkeit solcher Maßnahmen im bilingualen Unterricht liefern. Bisher sind jedoch keine Studien bekannt, die solche spezifischen methodisch-didaktischen Aspekte dieser Unterrichtsform in Bezug gesetzt haben zur Leistungsentwicklung der Klassen.

Abschließend scheint es, als ob einige der in den obigen Darstellungen identifizierten Forschungsdesiderate auf eine fehlende interdisziplinäre Zusammenarbeit verschiedener Wissenschaftsdomänen zurückgeführt werden könnten. Dies betrifft beispielsweise die Kooperation zwischen Fachdidaktiken und empirischer Bildungsforschung sowie die Zusammenarbeit zwischen Fremdsprachen- und Sachfachdidaktiken. Wurde bei bisherigen Analysen zum bilingualen Unterricht auf eine dieser Domänen verzichtet – so scheinen obig zitierte Studien zu demonstrieren –, dann schien oft eine Perspektive oder der korrekte forschungsmethodische Zugang zum Forschungsgegenstand zu fehlen. So wiesen Studien zur Wirksamkeit bilingualen Unterrichts, die überwiegend von Forscherinnen und Forschern der empirischen Bildungsforschung durchgeführt wurden, bislang einen sehr begrenzten und eher oberflächlichen fachdidaktischen Bezug auf, der sich meist auf Leistungstests und die Erfassung der Schulnoten in der Fremdsprache und den Sachfächern beschränkte (z. B. Köller et al., 2012; Nold et al., 2008). Spezifische methodisch-didaktische Aspekte, wie beispielsweise der Einsatz von Mutter- und Fremdsprache im bilingualen Unterricht und dessen mögliche Auswirkungen, blieben unberücksichtigt. Hierzu ist anzufügen, dass der bilinguale Sachfachunterricht in Deutschland grundsätzlich bislang eher wenig im Mittelpunkt des Forschungsinteresses der empirischen Bildungsforschung stand. Diese beschäftigte sich überwiegend mit einer anderen Ausprägung zweisprachigen Unterrichts, nämlich mit immersiven Lehr- und Lernformen (z. B. Baumert, Köller & Lehmann, 2012; Zaunbauer, Gebauer & Möller, 2012). Für fachdidaktisch geprägte Zugänge zur Erforschung der Leistungsentwicklung im bilingualen Unterricht ist festzustellen, dass meist nur die Fremdsprache (z. B. Rumlich, 2013; Fehling, 2008) oder, deutlich seltener, lediglich das Sachfach (z. B. Jäppinen, 2005) im Mittelpunkt des Forschungsinteresses standen. Studien hingegen, die die Wirksamkeit bilingualen Unterrichts sowohl hinsichtlich der Sprach- als auch der Sachfachkompetenzen der Lerner mithilfe von dafür notwendigen Längsschnittstudien untersuchten, sind kaum bekannt (als Ausnahme zu bilingualen Modulen z. B. Piesche et al., 2014). Abschließend scheinen einige fachdidaktische Studien zur Wirksamkeit bilingualen Unterrichts eine begrenzte Aussagekraft zu besitzen, da deren forschungsmethodischer Zugang Schwächen aufwies. Diesbezüglich wären beispielsweise

querschnittliche Studien zu nennen, die das Vorwissen der Schülerinnen und Schüler nicht berücksichtigen konnten (z. B. Zydatiß, 2007). Weiterhin wiesen manche Studien keine Vergleichsgruppen auf, anhand derer die Leistungen der bilingual unterrichteten Lerner verglichen und eingeordnet hätten werden können (z. B. D'Angelo, 2013; Abendroth-Timmer, 2007; Müller-Schneck, 2006). Zudem wurde in der Mehrzahl der Untersuchungen zur Wirksamkeit bilingualen Unterrichts der mögliche Effekt dieser Unterrichtsform auf die Kompetenzentwicklung der Lerner nicht von anderen Einflüssen bereinigt, die nicht in direktem Zusammenhang mit dieser Unterrichtsform standen, aber die Schülerleistung mit großer Wahrscheinlichkeit mitbeeinflussten. Dies betrifft Aspekte, wie beispielsweise individuelle bzw. klassenbezogene Lernvoraussetzungen (z. B. kognitive Grundfähigkeiten, sozioökonomischer Status, Motivation) sowie Lehrperson, Unterrichtsqualität und Leistungsemotionen. Ausnahmen bilden die DESI-Studie (DESI-Konsortium, 2008) sowie die Untersuchung von Rumlich (2014b), die zwar nicht alle, aber einige dieser Faktoren – überwiegend jedoch nur auf Schülerebene – statistisch kontrollierten.

1.7 Fragestellungen der vorliegenden Arbeit

1.7.1 Ableitung der Fragestellungen

Die vorliegende Arbeit beschäftigt sich mit der Wirksamkeit des bilingualen Unterrichts. In Anlehnung an Helmkes (2012) Angebots-Nutzungs-Modell wurden in den Abschnitten 1.2 bis 1.5 mehrere Aspekte aufgezeigt, die in diesem Kontext von besonderem Interesse zu sein scheinen. So wurde beispielsweise hinsichtlich der Aspekte „Familie“, „Lernpotential“ und „Klassenzusammensetzung“ festgestellt, dass sich bilingual unterrichtete Schülerinnen und Schüler bzw. Klassen aufgrund der möglichen Selektivität bilingualer Unterrichtsprogramme (vgl. Abschnitt 1.2.1) in ihren Lernvoraussetzungen von einsprachig unterrichteten Lernern bzw. Lernergruppen unterscheiden könnten, wie beispielsweise hinsichtlich des Vorkenntnisstands, kognitiver Grundfähigkeiten oder Motivation (vgl. Abschnitt 1.2.2). Dies konnte in Bezug auf die Schülerebene in einigen ersten Untersuchungen tentativ bestätigt werden (Rumlich, 2013 & 2014; Bos et al., 2009; Alonso et al., 2008; Fehling, 2008). Ähnliches trifft auf den Aspekt der „Lehrperson“ zu. Auch diesbezüglich verwiesen bisherige Überlegungen darauf, dass sich bilingual unterrichtende Lehrkräfte positiv von ihren einsprachig unterrichtenden Kolleginnen und Kollegen unterscheiden könnten, in Bereichen wie beispielsweise Fachkompetenz, schul- und unterrichtsbezogene Erwartungen, Ziele und Einstellungen sowie Motivation und Enthusiasmus (vgl. Abschnitt 1.3). Die Forschungsbasis hierzu ist recht dünn. Lediglich die Befunde von D'Angelo (2013) und Meyer (2003) deuteten an, dass bilingual unterrichtende Lehrkräfte besonders motiviert und motivierend in bilingualen Unterrichtsstunden zu sein scheinen. Weiterhin gaben die Überlegungen zum Aspekt „Unterrichtsangebot“ Grund zur Annahme, dass dieses im bilingualen Unterricht von höherer Qualität als im einsprachigen Sachfachunterricht sein könnte (vgl. Abschnitt 1.4.3). Dies kann mit

den möglicherweise positiveren Attributen bilingual unterrichtender Lehrkräfte im Vergleich zu ihren einsprachig unterrichtenden Kolleginnen und Kollegen begründet werden, welche in der Regel die Unterrichtsqualität beeinflussen (Kunter et al., 2013). Diesbezüglich ist lediglich die Untersuchung von Meyer (2003) bekannt, die andeutete, dass bilingual unterrichtete Schülerinnen und Schüler mehr Unterstützung durch die Lehrkraft, mehr schüleraktivierende Tätigkeiten sowie interessantere Lehrmethoden zu erfahren schienen. Schließlich scheint es auch bezüglich der „individuellen Wahrnehmung des Unterrichtsangebots“ möglich, dass sich diese zwischen den Lernern im bilingualen und einsprachigen Unterrichtssetting unterscheiden könnte. Dies trifft insbesondere auf die Leistungsemotionen der Schülerinnen und Schüler zu, da angenommen werden kann, dass bilingual unterrichtete Lerner, aufgrund ihrer grundsätzlich höheren Motivation und ihres gesteigerten Interesses am Unterricht, diesen auch positiver erleben (Pekrun et al., 2011; Götz et al., 2010; vgl. Abschnitt 1.4.4). Hierzu sind jedoch keine Studien bekannt, die auf den Kontext der vorliegenden Arbeit übertragbar wären. Die Betrachtung solcher möglichen Unterschiede zwischen bilingualem und einsprachigem Unterrichtssetting hinsichtlich Lerner, Lehrkräfte, Unterrichtsangebot und dessen Wahrnehmung ist jedoch nicht nur per se von Interesse. Für Untersuchungen zur leistungsbezogenen Wirksamkeit bilingualen Unterrichts müssen diese – im Idealfall auf Klassenebene – berücksichtigt und in Analysen konstant gehalten werden, da sie die Kompetenzentwicklung der Schülergruppen mitbeeinflussen. Somit soll der eigentliche Effekt dieser Unterrichtsform auf die Leistung der Klassen bestmöglich isoliert werden (vgl. Abschnitt 1.6.1).

Die Untersuchung genau dieses „Ertrags“ des bilingualen Unterrichts sowie der Vergleich mit dem des einsprachigen Sachfachunterrichts stellen das eigentliche Herzstück von Analysen zur Wirksamkeit bilingualen Unterrichts dar. Diesbezüglich geben die Überlegungen in den vorhergehenden Abschnitten Grund zur Annahme, dass sich hierbei Unterschiede zwischen den Schülergruppen zeigen könnten. Denn auf Basis verschiedener Theorien zum Zweitspracherwerb sowie weiterer Prinzipien, Modelle und Erkenntnissen aus der Fremdsprachen- und Geschichtsdidaktik sowie der Kognitionspsychologie kann argumentiert werden, dass bilingual unterrichtete Lerner bessere Leistungen in der Fremdsprache und im Sachfach erbringen könnten als einsprachig unterrichtete Schülerinnen und Schüler (vgl. Abschnitt 1.5). Dies betreffe sowohl die allgemeine englische Sprachfähigkeit und das englische Hörverstehen im Speziellen als auch das Sachwissen in Geschichte sowie die Einsicht der Lerner in den Konstruktionscharakter von Geschichte. Schülerkompetenzen im bilingualen Unterricht standen bisher immer wieder im Mittelpunkt des Forschungsinteresses; deutlich stärker etwa als die weiter oben genannten Aspekte, die die Schülerleistung beeinflussen und die es zu kontrollieren gilt. Bisherige Untersuchungen konnten zeigen, dass bilingual unterrichtete über deutliche bessere Englischkompetenzen zu verfügen scheinen als einsprachig unterrichtete Schülerinnen und Schüler (Rumlich, 2014b; Bos et al., 2009; Nold et al., 2008; Fehling, 2008; Bredenbröker, 2000). Die wenigen Befunde zum Sachfach deuten auf vergleichbare Leistungen zwischen den beiden Gruppen hin (Badertscher & Bieri, 2009;

Jäppinen, 2005). Bisherige Studien haben jedoch nur selten – wenn überhaupt – Aspekte wie Lernvoraussetzungen, Lehrkraft oder Unterrichtsqualität bei der Analyse der Leistungsentwicklungen von Schülerinnen und Schülern berücksichtigt. Weiterhin sind keine Studien bekannt, die die Kompetenzen ganzer bilingual unterrichteter Klassen in den Blick genommen und entsprechende Analysen auf Klassenebene durchgeführt haben.

Abschließend konnte in der vorliegenden Arbeit ein weiterer, fachspezifischer Aspekt aus dem Bereich des „Unterrichtsangebots“ identifiziert werden, der im Zusammenhang mit der Wirksamkeit bilingualen Unterrichts von besonderem Interesse zu sein scheint: Art und Umfang des Einsatzes von Mutter- und Fremdsprache. Dieser Aspekt stellt dasjenige Spezifikum dar, bezüglich dessen sich bilingualer Unterricht von anderen Unterrichtsformen unterscheidet (vgl. Abschnitt 1.4.1) und der daraufhin untersucht werden soll, wie sich die Verwendung der Sprachen im bilingualen Unterricht auf die Leistungen der Klassen auswirken könnte. Solche Analysen, die nur das bilinguale Unterrichtssetting betreffen, bedürfen demnach keiner Vergleiche mit einsprachigem Sachfachunterricht, sondern machen Aussagen zur Wirksamkeit verschiedener unterrichtlicher Maßnahmen in Bezug auf die Kompetenzentwicklung der bilingual unterrichteten Schülerklientel. Bisher sind keine Studien bekannt, die solche Untersuchungen durchgeführt haben. Deskriptive Befunde weisen darauf hin, dass Englisch die überwiegende Kommunikationssprache im bilingualen Unterricht darzustellen scheint und dass Deutsch vereinzelt, meist zweckbezogen zur Erreichung eines bestimmten Ziels, eingesetzt wird (Hüttner & Dalton-Puffer, 2013; Kollenrott, 2008; Gierlinger et al., 2007; Zydati, 2007; Müller-Schneck, 2006; Meyer, 2003; Bredenbröker, 2000).

1.7.2 Fragestellungen der empirischen Teilstudien

Die Überlegungen des vorhergehenden Abschnitts 1.7.1 zeigten zusammenfassend auf, dass zahlreiche Aspekte bestehen, die sich im Rahmen von Analysen zur Wirksamkeit bilingualen Unterrichts näher zu betrachten lohnen. Bisherige Forschungsarbeiten konnten hierzu bereits wertvolle Befunde liefern. Dennoch blieben bislang verschiedenste Zusammenhänge unerforscht (vgl. auch Abschnitt 1.6) und so scheinen nach wie vor viele Fragen offen, zu deren Beantwortung die folgenden empirischen Teilstudien beitragen wollen.

Teilstudie I.A: Lernausgangslagen von Schülerinnen und Schülern zu Beginn des bilingualen Geschichtsunterrichts. Unter Bezugnahme auf die beiden Aspekte „Familie“ und „Lernpotential“ von Helmkes (2012) Angebots-Nutzungs-Modell steht in Teilstudie I.A die Analyse der Lernausgangslagen bilingual und einsprachig unterrichteter Schülerinnen und Schüler im Mittelpunkt. Die Basis hierfür liefern Überlegungen zur möglichen Selektivität bilingualer Unterrichtsprogramme, die aufgrund der speziellen Anforderungen des bilingualen Unterrichts an seine Schülerinnen und Schüler (z. B. höhere kognitive Ansprüche, gesteigerter Lernaufwand) und den daraus resultierenden schulischen Auswahlverfahren bestehen könnte (vgl. Abschnitt 1.2.1). Demnach scheint es, als ob in Folge der spezifischen Organisation und Eigenschaften bilingualer

Unterrichtsprogramme soziale Selektionsprozesse stattfinden (Bonnet, 2012: 67), die dazu führen, dass meist besonders motivierte und leistungsstarke Kinder aus bildungsnahen Elternhäusern in bilinguale Züge aufgenommen werden (Bruton, 2011). Dies wurde in bisherigen Studien jedoch kaum explizit überprüft. Deshalb beschäftigt sich Teilstudie I.A mit der Frage, wie sich bilingual und einsprachig unterrichtete Schülerinnen und Schüler hinsichtlich ihrer Lernvoraussetzungen (verbale und figurale kognitive Grundfähigkeiten, sozioökonomischer Status, fachspezifische Motivation) voneinander unterscheiden und wie sich dies bezüglich ihrer Vorkenntnisse in den Fächern Englisch (allgemeine englische Sprachfähigkeit, Hörverstehen) und Geschichte (Sachwissen) verhält. Hierfür werden die Daten von $N = 1806$ Achtklässlern aus 75 Klassen herangezogen, die zu Schuljahresanfang, und somit zu Beginn des bilingualen Geschichtsunterrichts, erhoben wurden. Mithilfe von Mittelwertsvergleichen werden die Angaben von den 703 bilingual unterrichteten Lernern sowohl mit denen ihrer Parallelklassen als auch mit den Angaben von Schülerinnen und Schülern aus Schulen ohne bilinguales Unterrichtsangebot verglichen. Somit soll in Bezug auf die beiden Vergleichsgruppen noch zusätzlich untersucht werden, ob Nachteile für diejenigen Lerner entstehen könnten, die sich an Schulen mit bilingualem Zug nicht für dieses Profil, sondern für eine andere Spezialisierung der Schulen entschieden haben (vgl. Abschnitt 1.6.2).

Teilstudie I.B: Selection Effects in Content and Language Integrated Learning Classrooms in German Secondary Schools. In Teilstudie I.B werden mögliche Selektionseffekte bilingualer Unterrichtsprogramme untersucht. Sie knüpft an die Befunde der Teilstudie I.A an und beschäftigt sich auf Grundlage der dort gewonnenen Erkenntnisse zu Unterschieden in den Lernausgangslagen zwischen bilingual und einsprachig unterrichteten Schülerinnen und Schülern mit der Frage, ob sich die zu erwartenden Leistungsvorteile bilingual Unterrichteter in Englisch durch deren möglicherweise günstigere Lernvoraussetzungen erklären lassen. Hierzu wurden bisher kaum Untersuchungen durchgeführt. Deshalb wird in Teilstudie I.B in einem ersten Schritt die Frage untersucht, inwieweit sich die Schülergruppen hinsichtlich ihrer verbalen und figuralen Grundfähigkeiten, ihres familiären Hintergrunds (Bildungsabschluss der Eltern, sozioökonomischer Status, kulturelles Kapital) sowie ihrer Motivation und ihres bereits vorhandenen Sachwissens in Geschichte zu Beginn des bilingualen Geschichtsunterrichts in der 8. Klassenstufe voneinander unterscheiden. Im Unterschied zu Teilstudie I.A werden jedoch Mittelwertsvergleiche lediglich zwischen den 703 bilingual unterrichteten Lernern und den 659 Schülerinnen und Schülern aus den Parallelklassen ($N = 1362$) angestellt. Hiermit liegt der Fokus speziell auf Selektionseffekten *innerhalb* von Schulen mit bilingualem Zug. In einem zweiten Schritt wird die Frage beantwortet, inwieweit sich bilingual und einsprachig unterrichtete Lerner hinsichtlich ihrer allgemeinen englischen Sprachfähigkeit, dem englischen Hörverstehen und der Motivation in Englisch voneinander unterscheiden – diejenigen Aspekte also, die mit großer Wahrscheinlichkeit bereits durch die dreijährige Teilnahme der Lerner am bilingualen Zug beeinflusst worden sind. Im Anschluss daran bilden diese Unterschiedsanalysen die Basis zur Beantwortung der Frage, ob sich

die zu erwartenden Vorteile der bilingual unterrichteten Lerner in ihren Englischleistungen durch deren günstigere Lernvoraussetzungen in den Bereichen kognitive Grundfähigkeiten, familiärer Hintergrund, Sachwissen und Motivation in Geschichte anteilig erklären lassen; und, ob sich dieser Leistungsvorsprung nochmals reduziert, wenn zusätzlich Unterschiede in der Englischmotivation zwischen den beiden Schülergruppen berücksichtigt werden. Die Befunde beider Teilstudien I.A und I.B können als Basis für umfassende, längsschnittliche Untersuchungen zur Wirksamkeit bilingualen Unterrichts dienen.

Teilstudie II: The Effect of Content and Language Integrated Learning on Students' English and History Competences – Killing Two Birds With One Stone? Im Mittelpunkt der Analysen von Teilstudie II steht der „Ertrag“, also die Effekte des bilingualen Unterrichts auf die Kompetenzen der Schülerinnen und Schüler in Englisch und im Sachfach Geschichte. Dieser Forschungsfokus fußt auf Annahmen, wonach bilingual Unterrichtete bessere Leistungen erbringen könnten als einsprachig unterrichtete Klassen. Dies scheint auf Basis von Zweitspracherwerbstheorien und kognitionspsychologischen Erkenntnissen sowohl auf das Abschneiden in Englisch (Gass & Mackeys *interaction approach*, 2007; Krashen & Terrells *natural approach*, 2000; Krashens Input-Hypothese, 1985) als auch auf das Abschneiden in Geschichte zuzutreffen (Adesope et al., 2010; Cummins Schwellenhypothese, 1979; vgl. Abschnitt 1.5.1). Deshalb soll in Teilstudie II die Frage beantwortet werden, ob es Unterschiede in der Entwicklung der Englischleistungen (allgemeine englische Sprachfähigkeit, Hörverstehen) und des Sachwissens in Geschichte zwischen bilingual und einsprachig unterrichteten Klassen gibt unter Kontrolle verschiedener Aspekte, die möglicherweise zwischen den Klassen variieren (Vortestleistung, figurale und verbale kognitive Grundfähigkeiten, fachspezifisches Selbstkonzept, fachspezifischer intrinsischer Wert und Nützlichkeitswert, Geschlecht, Migrationshintergrund, sozioökonomischer Status, Unterrichtsqualität, Lehrermerkmale). Zur Beantwortung dieser Frage werden die Englisch- und Geschichtsleistungen von 30 bilingual und 45 einsprachig unterrichteten Klassen aus zwei verschiedenen Vergleichsgruppen ($N = 1806$ Lerner) miteinander verglichen. Die hierfür verwandten Mehrebenenregressionsanalysen basieren auf Daten, die am Anfang und Ende des (bilingualen) Geschichtsunterrichts der 8. Klassenstufe erhoben wurden. Dieses Forschungsvorhaben greift aktuelle Kritik zu bisherigen Untersuchungen zur Wirksamkeit bilingualen Unterrichts auf, wonach der Effekt dieser Unterrichtsform aufgrund querschnittlicher Untersuchungsdesigns (Rumlich, 2014a) und der damit in Verbindung stehenden, fehlenden Analyse der Lernvoraussetzungen von Schülerinnen und Schülern (Bruton, 2011) bisher oft überschätzt worden zu sein schien und es darüber hinaus grundsätzlich an Forschung zur Kompetenzentwicklung im Sachfach mangelt (Bonnet, 2012: 67; Coyle et al., 2010: 138). Mithilfe dieser Analysen können Aussagen zur Wirksamkeit bilingualen Unterrichts in Bezug auf die Kompetenzen der Lerner in der Fremdsprache sowie im Sachfach getroffen werden.

Teilstudie III: Merkmale erfolgreichen bilingualen Sachfachunterrichts – Die Rolle der Sprachen im deutsch-englischen Geschichtsunterricht. Teilstudie III beschäftigt sich in Anlehnung an den Aspekt des fachspezifischen „Unterrichtsangebots“ von Helmkes (2012) Angebots-Nutzungs-Modell mit dem Einsatz von Mutter- und Fremdsprache im bilingualen Geschichtsunterricht, was das zentrale Spezifikum dieser Unterrichtsform darstellt (vgl. Abschnitt 1.4.1). Obwohl Art und Umfang von Deutsch- und Englischesinsatz Teil der aktuellen fachdidaktischen Debatte ist (Diehr, 2012: 29) und dieser Aspekt in einigen Studien bereits untersucht wurde (z. B. Kollenrott, 2008; Zydati, 2007; Müller-Schneck, 2006), sind bisher keine Studien bekannt, die diese in Bezug gesetzt haben zu den Leistungen bilingual unterrichteter Klassen. Diesbezüglich deuten Theorien aus dem Bereich des Zweitspracherwerbs an (Gass & Mackey, 2007; Krashen, 1985; Cummins, 1979), dass die umfangreiche Verwendung von Englisch auf einem ausreichend hohen Niveau förderlich auf die Fremdsprachen- und Geschichtskompetenzen der Lerner im bilingualen Unterricht wirken könnte (vgl. Abschnitt 1.5.1). Weiterhin gibt die Übertragung von Butzkamms (1973) Prinzip der aufgeklärten Einsprachigkeit auf den Kontext des bilingualen Unterrichts Grund zur Annahme, dass auch der planvolle Einsatz der Muttersprache möglicherweise die Entwicklung der Schülerleistungen in Englisch sowie im Sachfach begünstigen könnte (vgl. Abschnitt 1.5.1). Auf dieser Grundlage möchte Teilstudie III zunächst die Frage beantworten, wie sich die Verwendung von Mutter- und Fremdsprache im bilingualen Geschichtsunterricht gestaltet, und zwar hinsichtlich des Umfangs von Englisch- und Deutscheinsatz durch Lehrkräfte und Lernende, des zweckgebundenen Einsatzes der Muttersprache zur Erreichung bestimmter Ziele sowie hinsichtlich der Sprachregel zur Einführung neuer Fachtermini. Die zweite Forschungsfrage zielt auf die Effekte ab, die diese sprachbezogenen methodisch-didaktischen Aspekte auf den Leistungszuwachs der Klassen in Englisch (allgemeine englische Sprachfähigkeit, Hörverstehen) und Geschichte (Sachwissen) zeigen könnten unter Kontrolle verschiedener Klassenmerkmale (sozioökonomischer Status, Vortestleistung, fachspezifischer intrinsischer Wert). Hierfür werden längsschnittlich erhobene Daten von 30 bilingual unterrichteten 8. Klassen ($N = 703$ Lerner) und deren bilingual unterrichtenden Geschichtslehrkräften zunächst deskriptiv und dann mehrbenenanalytisch ausgewertet. Entsprechende Befunde können Aussagen zur Wirksamkeit solcher spezifischer methodisch-didaktischer Maßnahmen hinsichtlich der Leistungsentwicklung von bilingual unterrichteten Klassen machen.

Vor Darstellung der empirischen Teilstudien in den Kapiteln 3 bis 6 erfolgt in Kapitel 2 zunächst eine detaillierte Erläuterung der konzeptionellen und forschungsmethodischen Ausrichtung der zugrunde liegenden Gesamtstudie „Competences and Motivation in Bilingual Instruction in History“ (COMBIH) mit Fokus auf deren interdisziplinär-triangularativen Gehalt. Jene Überlegungen wurden als Vorarbeit für die Analysen in den empirischen Teilstudien zu Beginn des COMBIH-Forschungsprojekts gesondert verfasst. Einschränkend ist diesbezüglich darauf hinzuweisen, dass

Kapitel 2 auf zwei Forschungsvorhaben verweist (Analyse der Leistungsemotionen und der Motivationsentwicklung der Schülerinnen und Schüler), welche noch nicht vollständig durchgeführt werden konnten. Dies lag einerseits an begrenzten zeitlichen Ressourcen sowie andererseits an einer leichten Verschiebung des Forschungsfokus hin zum Aspekt des Einsatzes von Mutter- und Fremdsprache im bilingualen Unterricht. Diese Umorientierung hatte sich erst im Laufe der Untersuchungen ergeben und war somit zurzeit der Anfertigung der Überlegungen des Kapitels 2 noch nicht bekannt. Weiterhin ist anzumerken, dass hinsichtlich der in den vorherigen Abschnitten 1.2 bis 1.5 als für den vorliegenden Kontext relevant identifizierten Aspekte die Unterschiedsanalysen bezüglich bilingual und einsprachig unterrichtenden Lehrkräften sowie die Analysen zur Entwicklung der geschichtsbezogenen Einsicht der Lerner in das Prinzip der Konstruktivität von Geschichte (vgl. Abschnitt 1.5.2) aus zeitlichen Gründen noch nicht abgeschlossen werden konnten. Abbildung 2 (nächste Seite) fasst in Anlehnung an Helmkes (2012) Angebots-Nutzungs-Modell nochmals zusammen, welche Aspekte in der vorliegenden Arbeit berücksichtigt und erhoben und welche davon in den empirischen Teilstudien analysiert wurden.

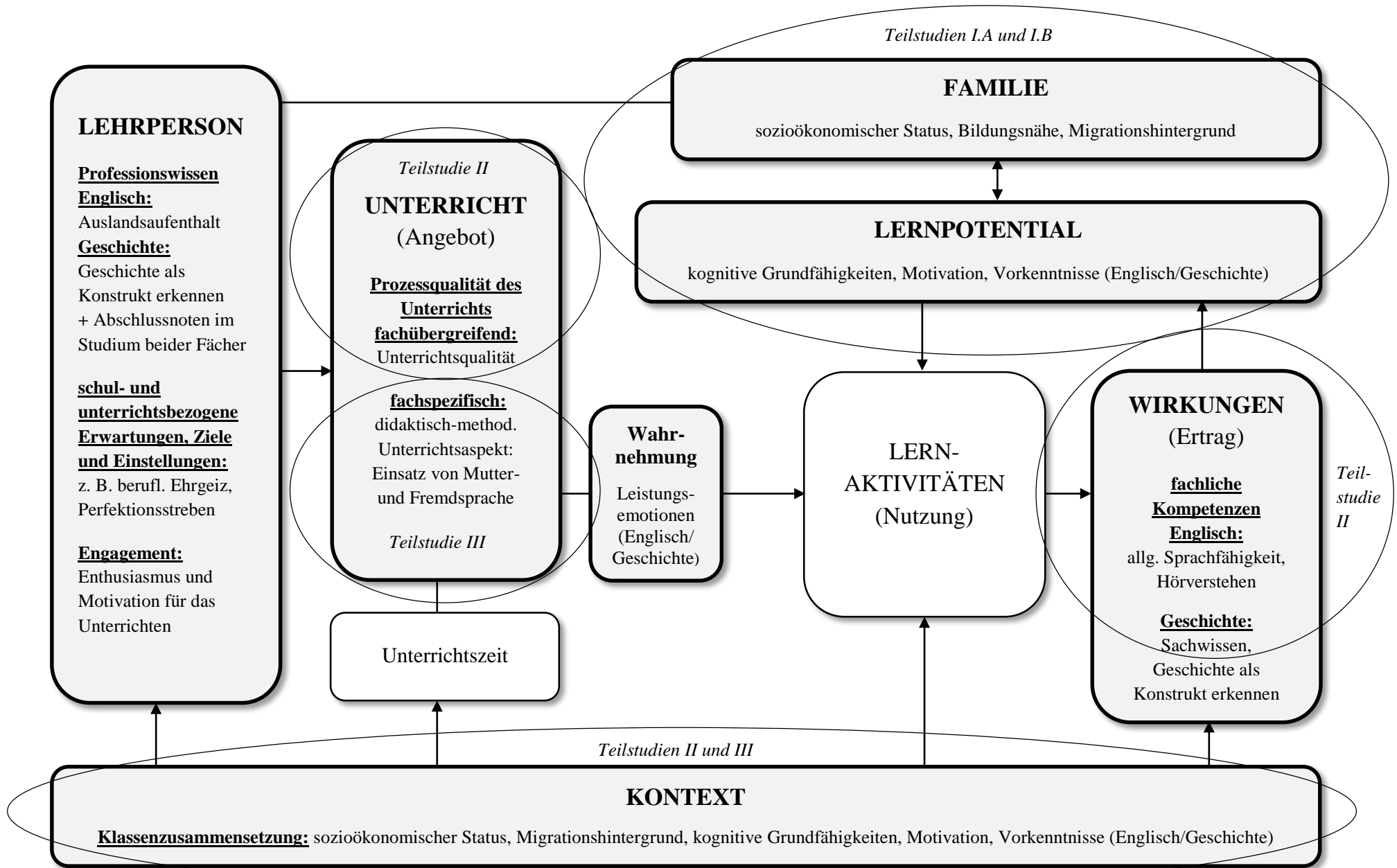


Abb. 2: Darstellung der auf Grundlage von Helmkes (2012) Angebots-Nutzungs-Modell in der vorliegenden Arbeit erhobenen (grau hinterlegt) sowie analysierten Aspekte (umkreist)

1.8 Literatur

- Abels, Heinz (2009). *Einführung in die Soziologie. Band 1: Der Blick auf die Gesellschaft*. Wiesbaden: VS Verlag.
- Abendroth-Timmer, Dagmar (2007). *Akzeptanz und Motivation: Empirische Ansätze zur Erforschung des unterrichtlichen Einsatzes von bilingualen und mehrsprachigen Modulen*. Frankfurt/ Main: Lang.
- Adesope, Olusola. O.; Lavin, Tracey; Thompson, Terri & Ungerleider, Charles (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research* 80(2), 207–245.
- Alonso, Esmeralda; Grisaleña, Jesús & Campo, Alejandro (2008). Plurilingual education in secondary schools: Analysis of results. *International CLIL Research Journal* 1(1), 36–49. URL: <http://www.icrj.eu/11/article3.html>, Zugriff am 08.12.2014.
- Bach, Gerhard & Timm, Johannes-Peter (Hrsg.) (2013). *Englischunterricht: Grundlagen und Methoden einer handlungsorientierten Unterrichtspraxis*. Stuttgart: UTB.
- Badertscher, Hans & Bieri, Thomas (2009). *Wissenserwerb im Content and Language Integrated Learning: Empirische Befunde und Interpretationen*. Bern: Haupt.
- Barricelli, Michele; Gautschi, Peter & Körber, Andreas (2012). Historische Kompetenzen und Kompetenzmodelle. In Michele Barricelli & Martin Lücke (Hrsg.), *Handbuch Praxis des Geschichtsunterrichts*. Schwalbach/ Ts.: Wochenschau Verlag, 207–235.
- Barricelli, Michele & Schmieder, Ulrich (2009). Über Nutzen und Nachteil des bilingualen Sachfachunterrichts. Fremdsprachen- und Geschichtsdidaktik im Dialog. In Daniela Caspari, Wolfgang Hallet, Anke Wegner & Wolfgang Zydati (Hrsg.), *Bilingualer Unterricht macht Schule. Beitrge aus der Praxisforschung*. Frankfurt/ Main: Lang, 205–220.
- Baumert, Jrgen; Kller, Olaf & Lehmann, Rainer (2012). Leseverstndnis im Englischen und Deutschen und Mathematikleistungen bilingual unterrichteter Schlerinnen und Schler am Ende der Grundschulzeit. *Unterrichtswissenschaft* 40, 290–314.
- Baumgartner, Hans-Michael (1997). Narrativitt. In Klaus Bergmann, Klaus Frhlich & Annette Kuhn (Hrsg.), *Handbuch der Geschichtsdidaktik*. Seelze: Kallmeyer, 157–160.
- Bernhardt, Markus (2011). Bilingualitt und historisches Lernen. Frderung von historischen Kompetenzen oder soziales Differenzkriterium? In Jan Hodel & Batrice Ziegler (Hrsg.), *Forschungswerkstatt Geschichtsdidaktik 09. Beitrge zur Tagung "Geschichtsdidaktik Empirisch 09" vom 3. und 4. September 2009 in Basel*. Bern: hep, 214–223.
- Bertram, Christiane; Wagner, Wolfram & Trautwein, Ulrich (2014). Zeitzeugenbefragungen im Geschichtsunterricht: Entwicklung eines Kurzinstruments fr die Wirksamkeitsmessung. In Tobias Arand & Manfred Seidenfu (Hrsg.), *Neue Wege - neue Themen - neue Methoden? Ein*

- Querschnitt aus der geschichtsdidaktischen Forschung des wissenschaftlichen Nachwuchses.* Göttingen: V&R unipress, 191–208.
- Bialystok, Ellen (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development* 70(3), 636–644.
- Bildungsserver Rheinland-Pfalz (1996 – 2014). *Gemeinschaftskunde: Sek. II (Bilingual – Englisch)*. URL: [http://lehrplaene.bildung-rp.de/lehrplaene-nachfaechern.html?tx_abdownloads_pi1\[action\]=](http://lehrplaene.bildung-rp.de/lehrplaene-nachfaechern.html?tx_abdownloads_pi1[action]=), Zugriff am 18.12.2014.
- Böing, Maik & Palmén, Paul (2012). Bilingual heißt zweisprachig! Überlegungen zur Verwendung beider Sprachen im bilingual deutsch-französischen Geographieunterricht. In Bärbel Diehr & Lars Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang, 73–90.
- Bonnet, Andreas (2012). Towards an evidence base for CLIL: How to integrate qualitative and quantitative as well as process, product and participant perspectives in CLIL research. *International CLIL Research Journal* 4(1), 66–78. URL: <http://www.icrj.eu/14/article7.html>, Zugriff am 08.12.2014.
- Bonnet, Andreas (2007). Fach, Sprache, Interaktion – Eine Drei-Säulen-Methodik für CLIL. *Fremdsprachen Lehren und Lernen* 36, 126–141.
- Bonnet, Andreas (2004). *Chemie im bilingualen Unterricht. Kompetenzerwerb durch Interaktion*. Opladen: Leske + Budrich.
- Borich, Gary D. (2007). *Observation skills for effective teaching*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Bos, Wilfried; Bosen, Martin & Gröhlich, Carola (2009). *KESS 7: Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*. Münster: Waxmann.
- Bredenbröcker, Winfried (2000). *Förderung der fremdsprachlichen Kompetenz durch bilingualen Unterricht. Empirische Untersuchungen*. Frankfurt/ Main: Lang.
- Breidbach, Stephan (2010). Bilinguale Didaktik – bald wieder zwischen allen Stühlen? Zu den Aussichten einer integrativen Didaktik des bilingualen Sachfachunterrichts. In Gerhard Bach & Susanne Niemeier (Hrsg.), *Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven*. Frankfurt/ Main: Lang, 165–176.
- Breidbach, Stephan (2007). *Bildung, Kultur, Wissenschaft. Reflexive Didaktik für den bilingualen Sachfachunterricht*. Münster: Waxmann.
- Breidbach, Stephan; Bach, Gerhard & Wolff, Dieter (Hrsg.) (2002). *Bilingualer Sachfachunterricht: Didaktik, Lehrer-, Lernerforschung und Bildungspolitik zwischen Theorie und Empirie*. Frankfurt/ Main: Lang.

- Breidbach, Stephan & Viebrock, Britta (2012). CLIL in Germany: Results from recent research in a contested field of education. *International CLIL Research Journal* 4(1), 5–16. URL: <http://www.icrj.eu/14/article1.html>, Zugriff am 08.12.2014.
- Bruton, Anthony (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System* 39(4), 523–532.
- Butzkamm, Wolfgang (2002). *Psycholinguistik des Fremdsprachenunterrichts. Von der Muttersprache zur Fremdsprache*. Tübingen: Francke.
- Butzkamm, Wolfgang (1973). *Aufgeklärte Einsprachigkeit. Zur Entdogmatisierung der Methode im Fremdsprachenunterricht*. Heidelberg: Quelle & Meyer.
- Cattell, Raymond B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology* 54, 1–22.
- Chomsky, Noam (1972). *Aspekte der Syntax-Theorie*. Frankfurt/ Main: Suhrkamp.
- Costa, Francesca & D'Angelo, Lauretta (2011). CLIL: A suit for all seasons. *Latin American Journal of Content & Language* 4(1), 1–13.
- Coyle, Do; Hood, Philip & Marsh, David (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Cummins, Jim (2006). *Language power and pedagogy. Bilingual children in the crossfire*. Clevedon: Multilingual Matters.
- Cummins, Jim (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49(2), 222–251.
- Dallinger, Sara (2013). Die Einstellung der Eltern zum bilingualen Sachfachunterricht: Eine quantitative Studie an Realschulen in Baden-Württemberg. In Jan Hollm, Armin Hüttermann, Jörg-U. Keßler, Gérald Schlemminger & Benjamin Ade-Thurow (Hrsg.), *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik, 268–286.
- Dalton-Puffer, Christiane; Nikula, Tarja & Smit, Ute (2010). Language use and language learning in CLIL. Current findings and contentious issues. In Christiane Dalton-Puffer, Tarja Nikula & Ute Smit (Hrsg.), *Language use and language learning in CLIL classrooms*. Amsterdam: Benjamins, 279–291.
- D'Angelo, Lauretta (2013). The construction of the CLIL subject teacher identity. In Stephan Breidbach & Britta Viebrock (Hrsg.), *Content and Language Integrated Learning (CLIL) in Europe. Research perspectives on policy and practice*. Frankfurt/ Main: Lang, 105–116.
- Daniel, Angelika (2001). *Lernerwortschatz und Wortschatzlernen im bilingualen Unterricht*. Frankfurt/ Main: Lang.
- Deci, Edward L. & Ryan, Richard M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik* 39(2), 223–238.

- DESI-Konsortium (Hrsg.) (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- DESI-Konsortium (Hrsg.) (2006). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistungen International (DESI)*. Frankfurt/ Main: DIPF. URL: <http://www.dipf.de/de/projekte/pdf/biqua/desi-zentrale-befunde>, Zugriff am 08.12.2014.
- Deutscher Bildungsserver (1996 – 2014). *Bildungspläne / Lehrpläne der Bundesländer für allgemeinbildende Schulen*. URL: <http://www.bildungsserver.de/Bildungsplaene-Lehrplaene-der-Bundeslaender-fuer-allgemeinbildende-Schulen-400.html>, Zugriff am 26.09.2014.
- Diederich, Jürgen & Tenorth, Heinz-Elmar (1997). *Theorie der Schule: Geschichte, Funktionen und Gestaltung*. Berlin: Cornelsen Scriptor.
- Diehr, Bärbel (2013). Evaluation, Leistungsmessung und Prüfungen im bilingualen Unterricht. In Wolfgang Hallet & Frank G. Königs (Hrsg.), *Handbuch Bilingualer Unterricht. Content and Language Integrated Learning*. Seelze: Kallmeyer, 209–215.
- Diehr, Bärbel (2012). What's in a name? Terminologische, typologische und programmatische Überlegungen zum Verhältnis der Sprachen im bilingualen Unterricht. In Bärbel Diehr & Lars Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang, 17–36.
- Diehr, Bärbel & Schmelter, Lars (Hrsg.) (2012). *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang.
- Dresel, Markus & Lämmle, Lena (2011). Motivation. In: Thomas Götz (Hrsg.), *Emotion, Motivation und selbstreguliertes Lernen*. Paderborn: Schöningh, 80–142.
- Duden (2013). Engagement. In Bibliographisches Institut GmbH (Hrsg.), *Duden – Die deutsche Rechtschreibung*. URL: <http://www.duden.de/rechtschreibung/Engagement>, Zugriff am 09.12.2014.
- Europarat (Hrsg.) (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Eurydice (Hrsg.) (2006). *Content and Language Integrated Learning (CLIL) at school in Europe*. URL: http://www.indire.it/lucabas/lkmw_file/eurydice/CLIL_EN.pdf, Zugriff am 09.12.2014.
- Fehling, Sylvia (2008). *Language Awareness und bilingualer Unterricht: Eine komparative Studie*. Frankfurt/ Main: Lang.
- Fend, Helmut (1981). *Theorie der Schule*. München: U&S Verlag.
- Frenzel, Anne C. & Stephens, Elizabeth J. (2011). Emotionen. In Thomas Götz (Hrsg.), *Emotion, Motivation und selbstreguliertes Lernen*. Paderborn: Schöningh, 15–77.

- Ganzeboom, Harry B.; de Graaf, Paul M.; Treimann, Donald J. & de Leeuw, Jan (1992). A standard international socio-economic index of occupational status. *Social Science Research* 25, 201–239.
- Gass, Susan M. & Mackey, Alison (2007). Input, interaction, and output in second language acquisition. In Bill VanPatten & Jessica Williams (Hrsg.), *Theories in second language acquisition. An introduction*. Mahwah, NJ: Erlbaum, 175–199.
- Gautschi, Peter (2009). *Guter Geschichtsunterricht: Grundlagen, Erkenntnisse, Hinweise*. Schwalbach/ Ts.: Wochenschau Verlag.
- Geiss, Peter (2013). Bilingualer Geschichtsunterricht. Ein Modell für das historische Lernen im *global village*. In Bärbel Diehr & Lars Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang, 55–71.
- Gierlinger, Erwin; Hametner, Christian & Spann, Harald (2007). Englisch als Arbeitssprache – viele Inseln und (noch) kein Festland? Eine Studie zum Arbeitssprachenunterricht auf Sekundarstufe I in Oberösterreich. In: Erwin Gierlinger, Verena Hainschink & Harald Spann (Hrsg.), *Content and Language Integrated Learning (CLIL) auf der Sekundarstufe I. Entwicklung, Forschung und Praxis*. Linz: Trauner, 63–118.
- Gottfredson, Linda J. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence* 24(1), 13–23. URL: <http://www.udel.edu/educ/gottfredson/reprints/1997mainstream.pdf>, Zugriff am 02.12.2014.
- Götz, Thomas; Cronjäger, Hanna; Frenzel, Anne C.; Lüdtke, Oliver & Hall, Nathan C. (2010). Academic self-concept and emotion relations: Domain specificity and age effects. *Contemporary Educational Psychology* 35, 44–58.
- Grotjahn, Rüdiger (2011). C-Tests: Aspekte der Validität. *Deutsch als Fremdsprache* 48(3), 131–137.
- Hallet, Wolfgang & Königs, Frank G. (Hrsg.) (2013). *Handbuch bilingualer Unterricht: Content and Language Integrated Learning*. Seelze: Kallmeyer.
- Hallet, Wolfgang & Königs, Frank G. (Hrsg.) (2010). *Handbuch Fremdsprachendidaktik*. Seelze: Kallmeyer.
- Hartmann, Ulrike (2008). *Perspektivenübernahme als eine Kompetenz historischen Verstehens*. URL: <https://ediss.uni-goettingen.de/handle/11858/00-1735-0000-0006-AD13-2>, Zugriff am 09.12.2014.
- Hartmannsgruber, Martin (2014). *Bilinguale Biologie: Konzeption und Evaluation*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Haß, Frank (Hrsg.) (2009). *Fachdidaktik Englisch: Tradition, Innovation, Praxis*. Stuttgart: Klett.
- Hasberg, Wolfgang (2004). Historisches Lernen im bilingualen Geschichtsunterricht(?). In Andreas Bonnet & Stephan Breidbach (Hrsg.), *Didaktiken im Dialog. Konzepte des Lehrens und Wege des Lernens im bilingualen Sachfachunterricht*. Frankfurt/ Main: Lang, 221–236.

- Hasberg, Wolfgang & Körber, Andreas (2003). Geschichtsbewusstsein dynamisch. In Andreas Körber (Hrsg.), *Geschichte – Leben – Lernen*. Schwalbach/ Ts.: Wochenschau Verlag, 177–200.
- Hattie, John (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Haugen, Einar (1956). *Bilingualism in the Americas: A bibliography and research guide*. Montgomery: University of Alabama Press.
- Heimes, Alexander (2013). Didaktiken und Methodiken bilingualer Fächer: Geschichte. In Wolfgang Hallet & Frank G. Königs (Hrsg.), *Handbuch Bilingualer Unterricht. Content and Language Integrated Learning*. Seelze: Kallmeyer, 345–352.
- Heine, Lena (2010). Fremdsprache und konzeptuelle Repräsentation: Bilingualer Unterricht aus kognitiver Perspektive. In Sabine Doff (Hrsg.), *Bilingualer Sachfachunterricht in der Sekundarstufe. Eine Einführung*. Tübingen: Narr, 199–212.
- Heil, Werner (2012). *Kompetenzorientierter Geschichtsunterricht*. Stuttgart: Kohlhammer.
- Helbig, Beate (2001). *Das bilinguale Sachfach Geschichte. Eine empirische Studie zur Arbeit mit französischsprachigen (Quellen-)Texten*. Tübingen: Stauffenburg.
- Heller, Kurt A. & Perleth, Christoph (2000). *Kognitiver Fähigkeitstest für 4. – 12. Klassen. Revision (KFT 4-12+R)*. Göttingen: Hogrefe.
- Helmke, Andreas (2012). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Kallmeyer.
- Helmke, Andreas & Klieme, Eckhard (2008). Unterricht und Entwicklung sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz, 301–312.
- Helmke, Andreas & Weinert, Franz E. (1997). Bedingungsfaktoren schulischer Leistung. In Franz E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule. Enzyklopädie der Psychologie: Pädagogische Psychologie*. Göttingen: Hogrefe, 71–176.
- Hermes, Liesel & Klippel, Friderike (Hrsg.) (2003). *Früher oder später? Englisch in der Grundschule und Bilingualer Sachfachunterricht*. München: Langenscheidt.
- Hollm, Jan (2013). Bilinguales Lehren und Lernen als schulische Innovation aus Sicht der Beteiligten. In Jan Hollm, Armin Hüttermann, Jörg-U. Keßler, Gérald Schlemminger & Benjamin Ade-Thurow (Hrsg.), *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik, 167–192.
- Hollm, Jan (2011). Wortschatzarbeit im bilingualen Sachfachunterricht. In Patrick Schäfer & Christine Schowalter (Hrsg.), *In mediam linguam. Mediensprache – Redewendungen – Sprachvermittlung*. Landau: Verlag Empirische Pädagogik, 319–329.
- Hollm, Jan; Hüttermann, Armin; Keßler, Jörg-U. & Schlemminger, Gérald (2010). BiliReal 2012: Bilinguale Züge für Englisch und Französisch in der Realschule. Erste Ergebnisse der

- wissenschaftlichen Begleitforschung zum Schulversuch in Baden-Württemberg. *Beiträge zur Fremdsprachenvermittlung* 49, 153–187. URL: http://www.vep-landau.de/fileadmin/user_upload/bzf/Hefte/bzf_2010_49.pdf, Zugriff am 02.12.2014.
- Hollm, Jan & Pirner, Manfred L. (2008). “The border is the place of understanding” (Paul Tillich). Religionsdidaktik und Englischdidaktik im Dialog. *Theo-Web. Zeitschrift für Religionspädagogik* 7(1), 105–123. URL: <http://www.theo-web.de/zeitschrift/ausgabe-2008-01/9.pdf>, Zugriff am 09.12.2014.
- Hüttner, Julia & Dalton-Puffer, Christiane (2013). Der Einfluss subjektiver Sprachlerntheorien auf den Erfolg der Implementierung von CLIL-Programmen. In Stephan Breidbach & Britta Viebrock (Hrsg.), *Content and Language Integrated Learning (CLIL) in Europe. Research perspectives on policy and practice*. Frankfurt/ Main: Lang, 129–144.
- ISB Bayern (Staatsinstitut für Schulqualität und Bildungsforschung München) (2006). *Kompetenz ...mehr als nur Wissen!* URL: <http://www.kompas.bayern.de/userfiles/infokompetenz.pdf>, Zugriff am 09.12.2014.
- Jäppinen, Aini-Kristiina (2005). Thinking and content learning of mathematics and science as cognitive development in Content and Language Integrated Learning (CLIL). Teaching through a foreign language in Finland. *Language and Education* 19(2), 148–169.
- Jeismann, Karl-Ernst (1980). „Geschichtsbewusstsein“. Überlegungen zur zentralen Kategorie eines neuen Ansatzes der Geschichtsdidaktik. In Hans Süsmuth (Hrsg.), *Geschichtsdidaktische Positionen: Bestandsaufnahme und Neuorientierung*, Paderborn: Schöningh, 179–222.
- Jespersen, Otto (1904). *How to teach a foreign language*. New York: Macmillan.
- Keßler, Jörg-U. & Schlemminger, Gérald (2013). Babylonisches Sprachengewirr: Wie benennen wir unseren Untersuchungsgegenstand? In Jan Holl, Armin Hüttermann, Jörg-U. Keßler, Gérald Schlemminger & Benjamin Ade-Thurow (Hrsg.), *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik, 15–26.
- Klein-Braley, Christine (1995). Leistungsmessung. In Karl-Richard Bausch, Herbert Christ & Hans-Jürgen Krumm (Hrsg.), *Handbuch Fremdsprachenunterricht*. Tübingen: Francke, 499–503.
- Klieme, Eckard; Avena, Hermann; Blum, Werner; Döbrich, Peter; Gruber, Hans et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: BMBF.
- Klieme, Eckard & Rakoczy, Katrin (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik* 54(2), 222–237.
- KM (Ministerium für Kultus, Jugend und Sport Baden-Württemberg) (Hrsg.) (2012). *Bilinguale Züge an Realschulen in Baden-Württemberg. Wissenswertes für Lehrkräfte*. URL: <http://www.kultusportal-bw.de/,Lde/Startseite/schulebw/Material+und+Links>, Zugriff am 22.08.2014.

- KM (Ministerium für Kultus, Jugend und Sport Baden-Württemberg) (Hrsg.) (2008). *Bilingualer Unterricht deutsch-englisch an allgemein bildenden Gymnasien*. Stuttgart: Schwäbische Druckerei.
- KM (Ministerium für Kultus, Jugend und Sport Baden-Württemberg) (Hrsg.) (2006). *Bilingualer Unterricht. Realschule*. Braunschweig: Westermann.
- KM (Ministerium für Kultus, Jugend und Sport Baden-Württemberg) (Hrsg.) (2004). *Bildungsplan 2004. Gymnasium*. URL: http://www.bildung-staerkt-menschen.de/service/downloads/Bildungsplaene/Gymnasium/Gymnasium_Bildungsplan_Gesamt.pdf, Zugriff am 30.09.2014.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Hrsg.) (2013). *Bericht „Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung“: Beschluss der Kultusministerkonferenz vom 17.10.2013*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/201_10_17-Konzepte-_bilingualer-_Unterricht.pdf, Zugriff am 09.12.2014.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Hrsg.) (2011). *Empfehlungen der Kultusministerkonferenz zur Stärkung der Fremdsprachenkompetenz. Beschluss der Kultusministerkonferenz vom 08.12.2011*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2011/2011_12_08-Fremdsprachenkompetenz.pdf, Zugriff am 09.12.2014.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Hrsg.) (2006). *Bericht „Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung“. Bericht des Schulausschusses vom 10.04.2006*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2006/2006_04_10-Konzepte-bilingualer-Unterricht.pdf, Zugriff am 09.12.2014.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Hrsg.) (2004). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-erste-Fremdsprache.pdf, Zugriff am 09.12.2014.
- Kollenrott, Anne (2008). *Sichtweisen auf deutsch-englisch bilingualen Geschichtsunterricht: Eine empirische Studie mit Fokus auf interkulturelles Lernen*. Frankfurt/ Main: Lang.
- Köller, Olaf; Leucht, Michael & Pant, Hans (2012). Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I. *Unterrichtswissenschaft* 4(4), 334–350.

- Königs, Frank G. (2013). Einsprachigkeit, Zweisprachigkeit und Code-Switching. In Wolfgang Hallet & Frank G. Königs (Hrsg.), *Handbuch Bilingualer Unterricht. Content and Language Integrated Learning*. Seelze: Kallmeyer, 174–180.
- Krashen, Stephen D. (1985). *The input hypothesis. Issues and implications*. London: Longman.
- Krashen, Stephen D. & Terrell, Tracey D. (2000). *The natural approach. Language acquisition in the classroom*. Harlow: Pearson Education.
- Krumm, Hans-Jürgen (2007). Lehr- Lernziele. In Karl-Richard Bausch, Herbert Christ & Hans-Jürgen Krumm (Hrsg.), *Handbuch Fremdsprachenunterricht*. Tübingen: Francke, 116–121.
- Kunter, Mareike; Baumert, Jürgen; Blum, Werner; Klusmann, Uta; Krauss, Stefan & Neubrand, Michael (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, Mareike; Klusmann, Uta; Baumert, Jürgen; Richter, Dirk; Voss, Thamar & Hachfeld, Axinja (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology* 105(3), 805–820.
- Küppers, Almut (2008). Bunte Berufsbiographien – ‚Referendariat‘ in England. Zur Situation von deutschen PGCE-Absolventen. In Susanne Ehrenreich, Gill Woodman & Marion Perrefort (Hrsg.), *Auslandsaufenthalte in Schule und Studium – Bestandsaufnahmen aus Forschung und Praxis*. Münster: Waxmann, 171–182.
- Lamsfuß-Schenk, Stefanie (2008). *Fremdverstehen im bilingualen Geschichtsunterricht. Eine Fallstudie*. Frankfurt/ Main: Lang.
- Long, Michael (1996). The role of the linguistic environment in second language acquisition. In Ritchie C. William & Bhatia K. Tej (Hrsg.), *Handbook of second language acquisition*. San Diego: Academic Press, 413–468.
- Long, Michael & Robinson, Peter (2005). Focus on form: Theory, research, and practice. In Catherine J. Doughty & Jessica Williams (Hrsg.), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press, 15–41.
- Mayer, Nikola & Köhler, Gudrun (2009). *Englischunterricht*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Mayer, Richard E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist* 59, 14–19.
- Meyer, Hilbert (2005). *Was ist guter Unterricht?* Berlin: Cornelsen Scriptor.
- Meyer, Christiane (2003). *Bedeutung, Wahrnehmung und Bewertung des bilingualen Geographieunterrichts. Studien zum zweisprachigen Erdkundeunterricht (Englisch) in Rheinland-Pfalz*. URL: <http://ubt.opus.hbz-nrw.de/volltexte/2004/192/pdf/20021118.pdf>, Zugriff am 09.12.2014.
- Müller-Bittner, Anke (2008). *Rezeption und Verwendung geographischer Lehr- und Lernmaterialien im bilingualen Sachfachunterricht. Interviewstudie mit Lehrkräften des deutsch-französischen*

- Erdkundeunterrichts in Nordrhein-Westfalen und Rheinland-Pfalz*. URL: <http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/MuellerBittnerAnke/diss.pdf>, Zugriff am 30.09.2014.
- Müller-Hartmann, Andreas & Schocker-von Ditfurth, Marita (2008). *Field trips* als Lernarrangements. Zur Ausbildung interkultureller Vermittlungskompetenzen bei angehenden Englischlehrkräften. In Susanne Ehrenreich, Gill Woodman & Marion Perrefort (Hrsg.), *Auslandsaufenthalte in Schule und Studium – Bestandsaufnahmen aus Forschung und Praxis*. Münster: Waxmann, 123–138.
- Müller-Schneck, Elke (2006). *Bilingualer Geschichtsunterricht. Theorie, Praxis, Perspektiven*. Frankfurt/ Main: Lang.
- Nold, Günter; Hartig, Johannes; Hinz, Silke & Rossa, Henning (2008). Klassen mit bilinguaem Sachfachunterricht: Englisch als Arbeitssprache. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz, 451–457.
- Oller, John W. & Hinofotis, Frances B. (1980). Two mutually exclusive hypotheses about second language ability: Indivisible or partially divisible competence. In John W. Oller & Kyle Perkins (Hrsg.), *Research in language testing*. Rowley, MA: Newbury House, 13–23.
- Otten, Edgar & Wildhage, Manfred (2007). Content and Language Integrated Learning. Eckpunkte einer „kleinen“ Didaktik des bilingualen Sachfachunterrichts. In Manfred Wildhage & Edgar Otten (Hrsg.), *Praxis des bilingualen Unterrichts*. Berlin: Cornelsen, 12–45.
- Pandel, Hans-Jürgen (2005). *Geschichtsunterricht nach Pisa: Kompetenzen, Bildungsstandards und Kerncurricula*. Schwalbach/ Ts.: Wochenschau Verlag.
- Pekrun, Reinhard (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review* 18, 315–341.
- Pekrun, Reinhard; Götz, Thomas; Frenzel, Anne C.; Barchfeld, Petra & Perry, Raymond (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology* 36, 36–48.
- Pérez-Cañado, María (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism* 15(3), 315–341.
- Piepho, Hans-Eberhard (1974). *Kommunikative Kompetenz als übergeordnetes Lernziel im Englischunterricht*. Dornberg-Frickhofen: Frankonius.
- Piesche, Nicole; Jonkmann, Kathrin & Keßler, Jörg-U. (2014). „Effekte des bilingualen Unterrichts auf das sachfachliche Lernen im naturwissenschaftlichen Unterricht - ein Feldexperiment“. Posterbeitrag auf der zweiten Tagung der Gesellschaft für empirische Bildungsforschung (GEBF) in Frankfurt am 03.03.2014.

- PISA-Konsortium (Hrsg.) (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Raudenbush, Stephen W. & Bryk, Anthony S. (2002). *Hierarchical linear models*. Thousand Oaks: Sage.
- Rautenhaus, Heike (2010). Prolegomena zu einer Didaktik des bilingualen Sachfachunterrichts. Beispiel: Geschichte. In Gerhard Bach & Susanne Niemeier (Hrsg.), *Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven*. Frankfurt/ Main: Lang, 109–120.
- Reinders, Heinz; Ditton, Hartmut; Gräsel, Cornelia & Gniewosz, Burkhard (2010). *Empirische Bildungsforschung: Gegenstandsbereiche*. Wiesbaden: VS Verlag.
- Rumlich, Dominik (2014a). Prospective CLIL and non-CLIL students' interest in English (classes): A quasi-experimental study on German sixth-graders. In Ruth Breeze, Concepción Pasamar, Carmen Saíz & Cristina Sala (Hrsg.), *Integration of theory and practice in CLIL*. Amsterdam: Rodopi, 75–95.
- Rumlich, Dominik (2014b). „Selektion und Präparation als Ursache für die hohen Sprachkompetenzen bilingual unterrichteter Lerner? Kritische Evaluation einer Längsschnittstudie“. Unveröffentlichter Vortrag auf der zweiten Tagung der Gesellschaft für empirische Bildungsforschung (GEBF) in Frankfurt am 05.03.2014.
- Rumlich, Dominik (2013). Students' general English proficiency prior to CLIL: Empirical evidence for substantial differences between prospective CLIL and non-CLIL students in Germany. In Stephan Breidbach & Britta Viebrock (Hrsg.), *Content and Language Integrated Learning (CLIL) in Europe: Research perspectives on policy and practice*. Frankfurt/ Main: Lang. 181–201.
- Rumlich, Dominik (2012). Anwendungsbeitrag. (Sprach-)Tests in der Praxis: Die Studie „Development of North Rhine-Westphalian CLIL Students“. In Sabine Doff (Hrsg.), *Fremdsprachenunterricht empirisch erforschen: Grundlagen – Methoden – Anwendung*. Tübingen: Narr, 169–178.
- Rüsen, Jörn (2008). *Historisches Lernen. Grundlagen und Paradigmen*. Schwalbach/ Ts.: Wochenschau Verlag.
- Rüsen, Jörn (1983). *Historische Vernunft. Grundzüge einer Historik I: Die Grundlagen der Geschichtswissenschaft*. Göttingen: Vandenhoeck Ruprecht.
- Schaarschmidt, Uwe & Fischer, Andreas W. (2008). *Arbeitsbezogenes Verhaltens- und Erlebensmuster (AVEM)*. Frankfurt/ Main: Pearson.
- Schönfeldt, Claudia (2005). *Die Rolle der Visualisierung im bilingualen deutsch-englischen Erdkundeunterricht*. URL: http://rzbl04.biblio.etc.tu-bs.de:8080/docportal/servlets/MCRFileNodeServlet/DocPortal_derivate_00000005/Document.pdf, Zugriff am 01.10.2014.

- Schreiber, Waltraud; Körber, Andreas; Borries van, Bodo et al. (2006). *Historisches Denken. Ein Kompetenz-Strukturmodell*. Neuried: ars una.
- Shulman, Lee S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher* 15(2), 4–14.
- Slavin, Robert E. (1997). *Educational psychology*. Boston: Allyn and Bacon.
- Springer Gabler (Hrsg.) (2014). Soziale Schicht. In Springer Gabler Verlag (Hrsg.), *Gabler Wirtschaftslexikon*. URL: <http://wirtschaftslexikon.gabler.de/Archiv/57703/soziale-schicht-v3.html>, Zugriff am 09.12.2014.
- Swain, Merrill (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Susan Gass & Carolyn Madden (Hrsg.), *Input in second language acquisition*. Rowley, MA: Newbury House, 235–256.
- Sylvén, Liss K. (2010). *Teaching in English or English teaching? On the effects of Content and Language Integrated Learning on Swedish learners' incidental vocabulary acquisition*. Göteborg: Acta Universitatis Gothoburgensis.
- Tej, Bhatia K. & William, Ritchie C. (2008). *The handbook of bilingualism*. Oxford: Wiley-Blackwell.
- Theis, Rolf (2010). Bilingualer Geschichtsunterricht. In Sabine Doff (Hrsg.), *Bilingualer Sachfachunterricht in der Sekundarstufe. Eine Einführung*. Tübingen: Narr, 44–57.
- Thürmann, Eike (2013). Spezifische Methoden für den bilingualen Unterricht/ CLIL. In Wolfgang Hallet & Frank G. Königs (Hrsg.), *Handbuch Bilingualer Unterricht. Content and Language Integrated Learning*. Seelze: Kallmeyer, 229–235.
- Thürmann, Eike (2010). Eine eigenständige Methodik für den bilingualen Sachfachunterricht? In Gerhard Bach & Susanne Niemeier (Hrsg.), *Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven*. Frankfurt/ Main: Lang, 71–89.
- Thurstone, Louis L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Timm, Johannes-Peter (2013). Schüleräußerungen und Lehrerfeedback im Unterrichtsgespräch. In Gerhard Bach & Johannes-Peter Timm (Hrsg.), *Englischunterricht: Grundlagen und Methoden einer handlungsorientierten Unterrichtspraxis*. Stuttgart: UTB, 199–229.
- Timm, Johannes-Peter (2005). *Englisch lernen und lehren: Didaktik des Englischunterrichts*. Berlin: Cornelsen.
- Viebrock, Britta (2007). *Bilingualer Erdkundeunterricht. Subjektive didaktische Theorien von Lehrerinnen und Lehrern*. Frankfurt/ Main: Lang.
- Vollmer, Helmut J. (2007). Leistungsmessung, Lernerfolgskontrolle, Selbstbeurteilung: Überblick. In Karl-Richard Bausch, Herbert Christ & Hans-Jürgen Krumm (Hrsg.), *Handbuch Fremdsprachenunterricht*. Tübingen: Francke, 365–370.
- Wagner, Rudi F. (2009). Lernen und Motivation. In Rudi F. Wagner, Arnold Hinz, Adly Rausch & Brigitte Becker (Hrsg.), *Modul Pädagogische Psychologie*. Stuttgart: UTB, 23–56.

- Weinert, Franz E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In Franz E. Weinert (Hrsg.), *Leistungsmessungen in Schulen*. Weinheim: Beltz, 17–31.
- Weskamp, Ralf (2013). Bildungsstandards lernerorientiert umsetzen. Kompetenzen, Aufgabenstellungen und Leistungsbeurteilung. In Gerhard Bach & Johannes-Peter Timm (Hrsg.), *Englischunterricht: Grundlagen und Methoden einer handlungsorientierten Unterrichtspraxis*. Tübingen: Francke, 256–279.
- Wigfield, Allan & Eccles, Jacquelynne (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology* 25, 68–81.
- Wild, Elke; Hofer, Manfred & Pekrun, Reinhard (2001). Lernmotivation. In Andreas Krapp & Bernd Weidenmann (Hrsg.), *Pädagogische Psychologie*. Weinheim: Beltz, 218–242.
- Wildhage, Manfred & Otten, Edgar (2003). *Praxis des bilingualen Unterrichts*. Berlin: Cornelsen Scriptor.
- Wolff, Dieter (1997). Bilingualer Sachfachunterricht: Versuch einer lernpsychologischen und fachdidaktischen Begründung. In Helmut J. Vollmer & Eike Thürmann (Hrsg.), *Englisch als Arbeitssprache im Fachunterricht: Begegnung zwischen Theorie und Praxis*. Soest: Gemeinsame Fachtagung der Deutschen Gesellschaft für Fremdsprachenforschung und des Landesinstituts für Schule und Weiterbildung, 50–62.
- Zaubauer, Anna; Gebauer, Sandra & Möller, Jens (2012). Englischleistungen immersiv unterrichteter Schülerinnen und Schüler. *Unterrichtswissenschaft* 40(4), 315–333.
- Zydati, Wolfgang (2007). *Deutsch-Englische Zge in Berlin (DEZIBEL): Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien*. Frankfurt/ Main: Lang.
- Zydati, Wolfgang (2005). *Bildungsstandards und Kompetenzniveaus im Englischunterricht: Konzepte, Empirie, Kritik und Konsequenzen*. Frankfurt/ Main: Lang.

Competences and Motivation in Bilingual Instruction in History (COMBIH)

– Eine Längsschnittstudie zu deutsch-englischem
Geschichtsunterricht an Gymnasien

Sara Dallinger

Abteilung Englisch, Pädagogische Hochschule Ludwigsburg

Kathrin Jonkmann

Abteilung Empirische Bildungsforschung und Pädagogische Psychologie, Universität Tübingen
Lehrgebiet Bildungspsychologie, FernUniversität Hagen

Dallinger, Sara & Jonkmann, Kathrin (2015). Competences and Motivation in Bilingual Instruction in History (COMBIH) – Eine Längsschnittstudie zu deutsch-englischem Geschichtsunterricht an Gymnasien. In Daniela Elsner & Britta Viebrock (Hrsg.), *Triangulation in der Fremdsprachenforschung*. Frankfurt/ Main: Lang, 171–189.
(Die Zitation folgt den Vorgaben des Sammelbands.)

Abstract

Die Studie COMBIH untersucht die Entwicklung der Englisch- und Geschichtsleistungen sowie der Motivation und Emotionen der Schülerinnen und Schüler im bilingualen Geschichtsunterricht. Dabei werden Eingangsunterschiede zwischen bilingual und regulär einsprachig unterrichteten Lernern zu Beginn des bilingualen Geschichtsunterrichts berücksichtigt. Der Schwerpunkt des Beitrags liegt auf der Darstellung und Erörterung des triangulativen Gehalts der quantitativen und längsschnittlichen Studie, bei der zu zwei Messzeitpunkten in Klassenstufe 8 Kompetenztests und Fragebögen für Schülerinnen und Schüler sowie Lehrkräfte an Gymnasien in Baden-Württemberg eingesetzt wurden.

Competences and Motivation in Bilingual Instruction in History (COMBIH) **– Eine Längsschnittstudie zu deutsch-englischem Geschichtsunterricht an Gymnasien**

Einleitung

In den letzten Jahren kann in einigen Bundesländern ein Trend zu einer verstärkten und flächendeckenden Einführung bilingualen Sachfachunterrichts beobachtet werden. Beispielsweise wurde in Baden-Württemberg an 65 Gymnasien ein bilingualer Zug eingerichtet und seit 2010 wird deutsch-englischer Sachfachunterricht auch an bis zu 20% aller Realschulen etabliert. In Nordrhein-Westfalen wurde das Programm „Bili für alle“ entwickelt (Breidbach/Viebrock 2012: 6). Diese politischen Entwicklungen stützen sich teilweise auf wissenschaftliche Studien, die zu überwiegend positiven Ergebnissen bezüglich des bilingualen Sachfachunterrichts kamen: Bilingual unterrichtete Schülerinnen und Schüler scheinen sowohl hinsichtlich ihrer Englisch- als auch ihrer Sachfachkompetenzen von dieser Unterrichtsform zu profitieren (Witzigmann 2011, Lamsfuß-Schenk 2008, Nold et al. 2008, Zydatiś 2007). Weiterhin zeigen sich sowohl Schulleitungen (Hollm et al. 2013), Lehrkräfte (Kollenrott 2008) als auch Eltern (Dallinger 2013) weitgehend überzeugt vom bilingualen Sachfachunterricht. Im Gegensatz zu dieser „Erfolgsgeschichte“ des bilingualen Lehrens und Lernens stehen Arbeiten, die diese in Frage stellen. Bruton (2011) verweist auf die in vielen Studien fehlende Analyse der Ausgangslage von bilingual und einsprachig unterrichteten Schülerinnen und Schülern, wonach bilingual unterrichtete Lerner meist schon zu Beginn des bilingualen Unterrichts die leistungstärkeren und motivierteren Schülerinnen und Schüler einer Schule sind. Ihre weiterhin positive Entwicklung könnte daher auf diese Eingangsunterschiede statt auf den bilingualen Unterricht zurückgeführt werden (Zydatiś 2012: 23). Dies führt vor allem bei querschnittlichen Studien zu einer systematischen Überschätzung des Effekts bilingualen Unterrichts (Rumlich 2014). So argumentiert auch Bonnet: „There are strong indications from evaluation research [...] that social selection plays an important role in German CLIL. Therefore, high outcomes in CLIL groups might well be ascribed to factors other than classroom practice“ (2012: 67). Es werden also Längsschnittstudien benötigt, die die unterschiedlichen Lernvoraussetzungen und Ausgangslagen von bilingual und einsprachig unterrichteten Schülerinnen und Schülern berücksichtigen und diese bei der Vorhersage von Veränderungen entsprechend kontrollieren, um somit den Effekt des bilingualen Unterrichts auf Schülerkompetenzen und Motivation bestmöglich isolieren zu können (Costa/D’Angelo 2011: 3). Die diesem Beitrag zugrunde liegende Studie COMBIH möchte versuchen diese Forschungslücken zu schließen.

Längsschnittliche Forschung zum bilingualen Sachfachunterricht

Die Unterschiede in den Lernausgangslagen bilingual und einsprachig unterrichteter Schülerinnen und Schüler gehen großteils auf die Auswahlprozesse zurück, die an Schulen mit fest eingerichtetem bilingualen Profil stattfinden. Diese Schulen besitzen neben anderen Spezialisierungen (wie z. B.

musisches oder naturwissenschaftliches Profil) einen bilingualen Zug, den Schülerinnen und Schüler bzw. Eltern am Ende der Grundschule wählen können. Die Auswahl seitens der Schulverwaltung basiert einerseits auf den Grundschulnoten und Empfehlungsschreiben der Grundschullehrkräfte. Andererseits stellen Gespräche zwischen Schulverwaltung und Eltern einen weiteren zentralen Baustein des Auswahlverfahrens dar (Zydati 2007: 70). Da in diesen Gesprchen hufig vermittelt wird, dass bilingualer Unterricht mit hheren kognitiven Ansprchen und Lernaufwand verbunden ist (Rumlich 2013a: 184, Hollm et al. 2010: 5), darf vermutet werden, dass sich vor allem besonders motivierte und interessierte Kinder mit hoher Lernbereitschaft fr den bilingualen Zug bewerben. In diesem Sinne kann von einer positiv selektierten Schlerschaft gesprochen werden, wenn diese bessere Lernvoraussetzungen aufweist als eine Vergleichsgruppe, wie etwa hhere kognitive Grundfhigkeiten (Helmke/Schrader 2001: 81), eine hhere Leistungsbereitschaft und Motivation (Wigfield/Eccles 2000) und eine grere Bildungsnhe des Elternhauses (Kller et al. 2012: 339). ber diese allgemeinen Lernvoraussetzungen hinaus ist es fr Studien, die erst im Laufe des bilingualen Zugs einsetzen, wichtig, die Unterschiede in den Englischleistungen und das Vorwissen im Sachfach zu erfassen und zu kontrollieren, um den inkrementellen Effekt des bilingualen Sachfachlernens im Laufe eines Schuljahres ermitteln zu knnen. Im Folgenden wird nun eine Auswahl an Studien prsentiert, die aufgrund ihres Lngsschnittdesigns diese Ausgangslagen der bilingual und einsprachig unterrichteten Schlerschaften bercksichtigen.

Mit C-Tests verglich Rumlich (2013a) in seiner Studie DENOCS die allgemeine Englischkompetenz einer Teilstichprobe von 340 bilingual und 430 nicht bilingual unterrichteten Gymnasialschlern. Die nicht bilingual unterrichtete Stichprobe setzte sich aus zwei Untergruppen zusammen: Schlerinnen und Schler aus den Parallelklassen der bilingual unterrichteten Klassen sowie Schlerinnen und Schler von Schulen ohne bilinguales Unterrichtsangebot. Die Ergebnisse der ersten Testung am Ende der Klasse 6, vor Beginn des bilingualen Sachfachunterrichts, verweisen auf deutlich hhere Englischleistungen der zuknftig bilingual unterrichteten Lerner. Bezglich der beiden Vergleichsgruppen schnitten die Schlerinnen und Schler aus Schulen ohne bilinguales Angebot besser ab als die Parallelklassenschler. Dies deutet darauf hin, dass innerhalb einer Schule eher die leistungsstrkeren Schlerinnen und Schler den bilingualen Zug whlen. Weiterhin wurde ein Fragebogen eingesetzt, der unter anderem das Schlerinteresse an der Fremdsprache und am Fach Englisch erfasste. Auch hier wurde festgestellt, dass bilingual unterrichtete Lerner mehr Interesse zeigten als die beiden nicht bilingual unterrichteten Schlergruppen (Rumlich 2014). Die Ergebnisse der zweiten Messung zum Ende der Klasse 7 verweisen darauf, dass der Leistungsvorsprung der bilingual unterrichteten Schlerinnen und Schler vom ersten Messzeitpunkt bezglich der allgemeinen Englischkompetenz zwar bestehen blieb, sich dieser aber im Verlauf eines Schuljahres nicht weiter vergrerte (Rumlich 2013b). Bilingual und einsprachig unterrichtete Lerner wiesen damit weitgehend parallel verlaufende Leistungszuwchse auf; bei den bilingual unterrichteten Schlerinnen und Schlern erfolgte dies lediglich auf einem hheren Niveau.

Fehling (2008) untersuchte in ihrer Studie die *Language Awareness*. Zum ersten Messzeitpunkt, zu Beginn Klasse 7, erhob sie Daten von 74 bilingual und 65 nicht bilingual unterrichteten Gymnasialschülern. Mithilfe des „*Language Awareness-Tests*“ und einem Test zur Erfassung der verbalen kognitiven Grundfähigkeiten konnte Fehling zeigen, dass bilingual unterrichtete Schülerinnen und Schüler in drei von vier gemessenen Dimensionen der *Language Awareness* (affektiv, kognitiv, politisch) bessere Resultate erzielten. Auch beim kognitiven Grundfähigkeitstest schnitten die bilingual unterrichteten Lerner besser ab. Im Verlauf der Klassenstufen 7 und 8 zeigten bilingual unterrichtete Schülerinnen und Schüler anschließend größere Zuwächse bezüglich der affektiven und kognitiven Dimension als einsprachig unterrichtete Lerner. Das Interesse an Englisch blieb bei den bilingual unterrichteten Schülerinnen und Schülern auf demselben Niveau wie zu Beginn des bilingualen Unterrichts. Bei den einsprachig unterrichteten Schülerinnen und Schülern nahm das Interesse an Englisch jedoch ab. Ein Absinken des Fachinteresses im Laufe der Sekundarstufe ist ein vielfach repliziertes entwicklungspsychologisches Phänomen (Krapp 2002, Wild/Hofer 2000). Das Entgegenwirken zu diesem Trend kann durchaus als Erfolg gewertet werden.

Zur Entwicklung von Kompetenzen im bilingual unterrichteten Sachfach liegen zurzeit kaum Studien vor (Bonnet 2012: 67, Coyle et al. 2010: 136-138). Badertscher/Bieri (2009) filmten in sechs Klassen einsprachig sowie bilingual unterrichtete Unterrichtseinheiten in den Fächern Geographie und Geschichte und führten vor und nach diesen Unterrichtsstunden Interviews mit den Schülerinnen und Schülern durch, um die Wissensstände zu erfassen. Dieses Vorgehen wurde nach 1,5 Jahren in den Klassen wiederholt. Das Forscherteam kam zu dem Schluss, dass keine Unterschiede zwischen den Schülergruppen bezüglich der Qualität des von den Lernern reproduzierten Sachfachwissens bestanden. Zu diesem Ergebnis kamen auch Alonso et al. (2008), die die Fremdsprachen- und Sachfachkompetenzen von 229 bilingual und einsprachig unterrichteten Sekundarstufenschülern im Baskenland untersuchten. Neben einem größeren Leistungszuwachs der bilingual unterrichteten Lerner in Bezug auf den Cambridge ESOL Test resümierte das Forscherteam, dass diese Schülergruppe ähnliche Leistungen wie die einsprachig unterrichteten Lerner in den Sachfächern zeigte. Diese Aussage stützte sich jedoch lediglich auf die subjektive Einschätzung von Lehrkräften sowie auf die Noten der Schülerinnen und Schüler, die Referenzgruppeneffekten unterliegen (Baumert et al. 2003). Zusammenfassend gilt anzumerken, dass in diesen Studien keine standardisierten und objektiven Leistungstests zur Erfassung des Sachfachwissens eingesetzt wurden.

Auch in Bezug auf die Effekte bilingualen Unterrichts auf die Veränderung der Motivation im Sachfach und in der Fremdsprache ist die empirische Datenbasis recht klein (Lasagabaster/Sierra 2009: 4). Abendroth-Timmer (2004) untersuchte bilingual und mehrsprachig unterrichtete Module der Sekundarstufe in verschiedenen Sachfächern für die Arbeitssprachen Englisch, Französisch und Spanisch. Mithilfe von Einzelfallanalysen, die auf Schülerinterviews basierten, und den Daten von

Schülerfragebögen kam sie zu dem Schluss, dass bilingualer Unterricht für die Motivation der Schülerinnen und Schüler nicht abträglich sei, dies jedoch von verschiedenen Faktoren abhängen. So kann bilingualer Unterricht bei angemessenem Sprach- und inhaltlichem Niveau sowie unter Berücksichtigung sachfachdidaktischer Prinzipien seitens der Lehrkraft (z. B. Authentizität, Handlungsbezug) förderlich auf die Motivation wirken. Zahlreiche querschnittliche Studien bestätigen dieses Ergebnis (z. B. Lamsfuß-Schenk 2008, Zydatiß 2007). Diese Befunde sollten jedoch mittels standardisierter Fragebögen zur differenzierten Erfassung unterschiedlicher motivationaler Konstrukte wie z. B. Fähigkeitsüberzeugungen, Fachinteresse, Nutzenbewertung oder Leistungsemotionen im Rahmen längsschnittlicher Studien auf eine breitere empirische Basis gestellt werden.

Abschließend werden zwei Studien querschnittlichen Designs vorgestellt, die zwar nicht die Ausgangslagen in Englisch, dafür aber zentrale Indikatoren der Lernvoraussetzungen bilingual und einsprachig unterrichteter Schülerinnen und Schüler berücksichtigten. Nold et al. (2008) untersuchten mithilfe einer Teilstichprobe von DESI die Englischkompetenzen von rund 2000 Gymnasialschülern. Hierbei wurden die Daten des zweiten Messzeitpunkts (Ende der 9. Klassenstufe) herangezogen. Um dem Problem der positiveren Ausgangslage bilingual unterrichteter Lerner zu begegnen, wurden die Schülerinnen und Schüler der Vergleichsgruppe gezielt ausgewählt. Es wurden 987 Lerner bestimmt, die sich auf Basis der Daten des ersten Messzeitpunkts (Beginn der 9. Klassenstufe) nicht von den bilingual unterrichteten Schülerinnen und Schülern unterschieden hinsichtlich ihres Geschlechts, der Erstsprache, Schulabschluss und Beruf der Eltern, der Note in Deutsch und den kognitiven Grundfähigkeiten. Die Ergebnisse zeigten, dass bilingual unterrichtete Schülerinnen und Schüler in vier Kompetenzbereichen (Hör- und Leseverstehen, Grammatik, Textrekonstruktion) das jeweils höhere Kompetenzniveau in Bezug auf den Gemeinsamen Europäischen Referenzrahmen für Sprachen erreichten als die einsprachig unterrichteten Lerner. Die größten Leistungsvorteile traten im Bereich des Hörverstehens auf (DESI-Konsortium 2006: 59). Köller et al. (2012) konnten diese Befunde replizieren. Sie analysierten die Daten einer Teilstichprobe von 345 bilingual und 9521 einsprachig unterrichteten Schülerinnen und Schülern der 9. Klassenstufe bezüglich ihrer Englischkompetenzen (Hör- und Leseverstehen). Folgende Hintergrundmerkmale wurden statistisch kontrolliert, um die Unterschiedlichkeit der beiden Schülergruppen zu berücksichtigen: Geschlecht, Erstsprache, Migrationshintergrund, sozioökonomischer Hintergrund, Beruf der Eltern, Deutschkompetenzen der Lerner. Trotz Kontrolle dieser Einflüsse auf die Englischleistungen zeigte sich, dass die bilingual unterrichteten Schülerinnen und Schüler deutlich besser in den Tests abschnitten als die regulär einsprachig unterrichteten Lerner. Zusammenfassend scheinen bilingual unterrichtete Schülerinnen und Schüler eine Schülergruppe zu sein, die über besonders gute Lernvoraussetzungen verfügt. Diese Unterschiede müssen Studien, die die Effekte bilingualen Unterrichts bestimmen möchten, kontrollieren. Empirische Befunde weisen beispielsweise darauf hin, dass der Leistungsvorsprung in Englisch auch dann noch besteht, wenn

zentrale Hintergrundmerkmale (z. B. kognitive Grundfähigkeiten, Bildungsnähe des Elternhauses) kontrolliert werden. Ob diese Unterschiede jedoch auf den erhöhten Stundenumfang in Englisch in den Klassenstufen 5 und 6, der von Schülerinnen und Schülern eines bilingualen Zugs in der Regel geleistet werden muss, zurückgehen, lassen die Daten der zitierten Studien offen, da die erste Messung zwar immer zu Beginn des bilingualen Sachfachunterrichts, jedoch nicht schon zu Beginn des verstärkten Englischunterrichts durchgeführt wurde. Weiterhin deuten Rumlichs Ergebnisse an (2013b), dass durch den bilingualen Sachfachunterricht keine inkrementellen Zuwächse erzielt werden, während Fehling (2008) für Englisch größere Leistungszuwächse bilingual unterrichteter Lerner feststellt. Auch in Bezug auf die Entwicklung von Sachfachkompetenzen sowie von Motivation und Emotionen liegen bisher nur wenig und zum Teil widersprüchliche Ergebnisse vor. Es bedarf an weiterer, belastbarer empirischer Forschung zur Validierung und Ausdifferenzierung dieser Befunde, die nicht nur zentrale Aspekte wie Kompetenzen und Motivation von Schülerinnen und Schülern im bilingualen Unterricht in Blick nimmt, sondern auch wichtige Kontextvariablen und Hintergrundmerkmale berücksichtigt (Pérez-Cañado 2012: 330-332).

Forschungsfragen

Für die COMBIH-Studie ergibt sich aus den dargestellten Befunden folgender Forschungsschwerpunkt: Wie wirkt sich bilingualer Geschichtsunterricht auf die Leistungsentwicklung, Motivation und Emotionen der Schülerinnen und Schüler in den Fächern Englisch und Geschichte aus? Zur Beantwortung dieser Frage wird ein längsschnittliches Untersuchungsdesign mit einem ersten Messzeitpunkt zu Beginn des bilingualen Geschichtsunterrichts in Klassenstufe 8 durchgeführt, um Unterschiede in den Ausgangslagen zwischen den Schülerschaften und gegebenenfalls zwischen den Lehrerschaften kontrollieren und somit potentielle differentielle Veränderungen im Laufe der Klasse 8 bestmöglich auf den bilingualen Unterricht zurückführen zu können. Selbstverständlich ist auch bei diesem Untersuchungsdesign die Aussagekraft über Ursache-Wirkungs-Zusammenhänge eingeschränkt und davon abhängig, ob die zentralen konfundierenden Variablen gemessen und kontrolliert werden. Echte Kausalaussagen sind nur mit einem experimentellen Design und randomisierter Bedingungszuweisung möglich, bei dem Störvariablen und Alternativerklärungen ausgeschlossen werden können. Da dies in Bildungssettings meist nicht realisierbar ist, muss mit Längsschnittstudien und der Verwendung geeigneter statistischer Verfahren eine bestmögliche Approximation geleistet werden. Die Forschungsfragen von COMBIH lauten wie folgt:

Lernvoraussetzungen und Ausgangslagen von bilingual und einsprachig unterrichteten Schülerinnen und Schülern zu Beginn der 8. Klassenstufe

- Inwiefern unterscheiden sich bilingual und einsprachig unterrichtete Schülerinnen und Schüler zu Beginn des bilingualen Geschichtsunterrichts hinsichtlich ihres familiären Hintergrunds, ihrer

kognitiven Grundfähigkeiten, ihrer Schulnoten sowie ihrer Leistungen, Motivation und Leistungsemotionen in den Fächern Englisch und Geschichte?

- Unterscheiden sich die Englisch- und Geschichtsleistungen von bilingual und einsprachig unterrichteten Schülerinnen und Schülern auch dann noch, wenn zentrale Hintergrundmerkmale kontrolliert werden?
- Gibt es Unterschiede zwischen den beiden nicht bilingual unterrichteten Vergleichsgruppen hinsichtlich der zentralen Konstrukte?

In Anlehnung an die berichteten Ergebnisse zu den Ausgangslagen der beiden Schülergruppen wird erwartet, dass die bilingual unterrichteten Schülerinnen und Schüler in den Leistungstests und Schulnoten besser abschneiden als die einsprachig unterrichteten Lerner. Außerdem werden für beide Fächer positivere Motivationswerte und Emotionen bei den bilingual Unterrichteten erwartet. Darüber hinaus wird angenommen, dass der Leistungsvorsprung der bilingualen Lerner in Englisch auch dann noch besteht, wenn Hintergrundmerkmale kontrolliert werden. Die Leistungsvorsprünge in Geschichte hingegen sollten weitgehend durch die Unterschiede in den kognitiven Grundfähigkeiten und im familiären Hintergrund erklärt werden. Weiterhin wird angenommen, dass die Schülerinnen und Schüler der nicht bilingual unterrichteten Parallelklassen niedrigere Leistungen und Motivation aufweisen als die Schülerinnen und Schüler von Schulen ohne bilinguales Unterrichtsangebot.

Veränderungen von Kompetenzen, Motivation und Emotionen in der 8. Klassenstufe in Abhängigkeit vom bilingualen Unterricht

- Können bilingual unterrichtete Schülerinnen und Schüler einen möglichen Leistungsvorsprung gegenüber einsprachig unterrichteten Schülerinnen und Schülern in den Fächern Englisch und Geschichte im Lauf eines Jahres bilingualen Geschichtsunterrichts weiter ausbauen?
- Verändern sich die Motivation und die Leistungsemotionen der bilingual unterrichteten Schülerinnen und Schüler in den Fächern Englisch und Geschichte positiver als bei einsprachig unterrichteten Schülerinnen und Schülern?
- Bestehen diese differentiellen Verläufe auch dann noch, wenn die Ausgangslagen, die Hintergrundmerkmale und Kontextvariablen kontrolliert werden?

Bezüglich der Englischkompetenzen wird erwartet, dass die bilingual unterrichteten Schülerinnen und Schüler einen größeren Leistungszuwachs erfahren als die einsprachig unterrichteten Lerner, und dass diese Unterschiede auch nach Kontrolle der Ausgangslagen, Hintergrundmerkmale und Kontextvariablen statistisch bedeutsam bleiben. Wie sich die Kompetenzen im Fach Geschichte entwickeln, ist eine offene Forschungsfrage. Bezüglich der Motivation und Emotionen bilingual unterrichteter Schülerinnen und Schüler kann erwartet werden, dass diese auf demselben (erhöhten) Niveau bleiben wie beim ersten Messzeitpunkt, während für die einsprachig unterrichteten Schülerinnen und Schüler der typische Entwicklungsverlauf eines sinkenden Interesses vorliegt.

Design der Studie

Stichprobe

Die Stichprobe der Studie bestand aus Schülerinnen und Schülern der 8. Klassenstufe von Gymnasien in Baden-Württemberg. Sie gliederte sich in drei Untersuchungsgruppen auf. Die erste Gruppe setzt sich aus Schülerinnen und Schülern zusammen, die ab Klasse 5 den bilingualen Zug ihrer Schule besuchten. In den Klassenstufen 5 und 6 nahmen sie am verstärkten Englischunterricht teil (2 zusätzliche Schulstunden pro Woche) und ab Klasse 7 wurden sie in mindestens einem Fach pro Schuljahr durchgängig bilingual unterrichtet. In Klasse 7 war dies Geographie, in der 8. Klasse nahmen die Schülerinnen und Schüler zum ersten Mal an bilinguaalem Geschichtsunterricht teil. Regulär deutschsprachiger Geschichtsunterricht wurde bereits ab Klasse 6 erteilt. Die Schülerinnen und Schüler der Vergleichsgruppe I besuchten die Parallelklassen der bilingualen Gruppen. D. h., sie hatten sich gegen den bilingualen Zug entschieden oder wurden in diesen nicht aufgenommen. Vergleichsgruppe II umfasste Schülerinnen und Schüler von Gymnasien ohne bilinguales Unterrichtsangebot. Die Lehrerstichprobe setzte sich aus den Geschichtslehrkräften der an der Studie beteiligten Klassen zusammen. Um eine hinreichend hohe statistische Teststärke zu gewährleisten, war es das Ziel der Studie, mindestens 20 Klassen pro Gruppe zu rekrutieren. Um dies zu erreichen, wurden alle 65 Gymnasien mit bilinguaalem Zug in Baden-Württemberg gezielt kontaktiert und um Teilnahme gebeten. Die Schulen der Vergleichsgruppe II wurden aus Machbarkeitsgründen aus dem näheren Umkreis der Hochschulen gewählt. 28 der Gymnasien mit bilinguaalem Zug nahmen an der Studie teil sowie 9 weitere Gymnasien ohne bilinguales Angebot (vgl. Tab. 1). An der Vorstudie, die der Pilotierung von Teilen der Erhebungsmaterialien diente, nahmen 12 Klassen aus vier Gymnasien mit 222 Schülerinnen und Schülern der Klassenstufen 7 und 8 teil.

Tab. 1: Stichprobe der Hauptstudie

| Gymnasien | | Klassen | | Schüler MZP1 | Schüler MZP2 |
|--------------------------|-----------|---------------------|-----------|--------------|--------------|
| mit bilinguaalem Zug | 28 | Bilingual-Klassen | 30 | 664 | 661 |
| | | Vergleichsgruppe I | 28 | 588 | 577 |
| ohne bilinguales Angebot | 9 | Vergleichsgruppe II | 17 | 411 | 394 |
| Gesamt | 37 | | 75 | 1663 | 1632 |

Anmerkung: MZP = Messzeitpunkt.

Organisation und Ablauf der Studie

Vor Beginn der Datenerhebungen wurden alle eingesetzten Materialien und Prozeduren für eine datenschutzrechtliche Prüfung an das Ministerium für Kultus, Jugend und Sport Baden-Württemberg geleitet. Diese wurden dort begutachtet und für die Untersuchung freigegeben. Der erste Messzeitpunkt fand im September und Oktober 2012 statt. Dieser Zeitpunkt wurde gewählt, um die Schülerinnen und Schüler und Lehrkräfte zu Beginn des Schuljahres befragen zu können. Dieselbe

Schüler- und Lehrerstichprobe wurde im Juni und Juli 2013, zum Ende des Schuljahres, ein zweites Mal befragt. Zum ersten Messzeitpunkt nahm die Datenerhebung insgesamt 135 Minuten in Anspruch (exklusive der Pausenzeiten), zur zweiten Messung 115 Minuten. Die Erhebungen wurden mithilfe eines standardisierten Manuals von geschulten Testleitern durchgeführt. Zusätzlich waren Lehrkräfte im Klassenzimmer anwesend. Das Test- und Befragungsmaterial der Schülerinnen und Schüler wurde vor Ort ausgeteilt und nach Beendigung der Erhebung wieder von den Testleitern eingesammelt. Die Lehrerfragebögen, die ca. 20 Minuten in Anspruch nahmen, wurden entweder an der Schule oder zuhause ausgefüllt.

Erhebungsinstrumente

Die Untersuchungsmaterialien der Schülerinnen und Schüler umfassten Leistungstests und einen Fragebogen. Die Englischleistungen wurden mittels eines eigens entwickelten C-Tests sowie mit einem Hörverstehenstest aus dem Aufgabenpool der Bildungsstandards (IQB) geprüft. Das Geschichtswissen der Schülerinnen und Schüler aus den Klassen 7 und 8 wurde mithilfe von Lückentexten und *Multiple Choice*-Aufgaben gemessen, die in Kooperation mit Geschichtsdidaktikern entwickelt wurden. Außerdem wurden Fragen zur Erfassung der historischen De- und Rekonstruktionskompetenz (Bertram 2012) in Anlehnung an das Kompetenz-Strukturmodell für historisches Denken nach Schreiber et al. (2006) integriert. Die figuralen und verbalen kognitiven Grundfähigkeiten der Schülerinnen und Schüler wurden mit Subtests des Kognitiven Fähigkeitstests erhoben (KFT; Heller/Perleth 2000). Die Fragebogenitems zur Schülermotivation in den Fächern Englisch und Geschichte orientierten sich an dem Erwartungs-Mal-Wert-Modell nach Wigfield/Eccles (2000), das sich bereits in vielen pädagogisch psychologischen Studien bewährt hat. In diesem Modell werden unterschiedliche motivationale Konstrukte zur Erklärung leistungsbezogenen Verhaltens berücksichtigt, wie beispielsweise Interesse, subjektiver Nutzen und akademisches Selbstkonzept. Die Erfassung fachspezifischer Leistungsemotionen (z. B. Freude, Langeweile) erfolgte auf Basis der Kontroll-Wert-Theorie nach Pekrun (2006, Götz et al. 2010, Götz/Frenzel 2010). Im Schülerfragebogen befanden sich weiterhin Fragen zum bilingualen Geschichtsunterricht (z. B. Bewertung der Unterrichtsform, Rolle des Englischen) sowie allgemeine Fragen zur Unterrichtsqualität in den Fächern Englisch und Geschichte (COACTIV-Studie, Kunter et al. 2011). Die Operationalisierung des familiären Hintergrunds orientierte sich an den PISA-Studien. Außerdem wurden die Schulnoten der Lerner erhoben. Die Geschichtslehrkräfte der betreffenden Klassen erhielten einen Fragebogen mit eigens entwickelten Fragen zu methodisch-didaktischen Aspekten des (bilingualen) Geschichtsunterrichts, ihrer Lehrerbiographie sowie zu arbeitsbezogenen Verhaltens- und Erlebensmustern (AVEM; Schaarschmidt/Fischer 2008).

Datenanalyse

Die Studie wird derzeit unter Verwendung der Softwarepakete SPSS und Mplus insbesondere mit regressionsanalytischen Verfahren ausgewertet. Da ganze Klassen in die Untersuchung aufgenommen wurden, liegen keine vollständig unabhängigen Beobachtungen vor. Diese geschachtelte Datenstruktur wird bei den Analysen berücksichtigt. Andernfalls würden die Standardfehler unterschätzt und es erfolgt eine zu liberale statistische Signifikanzprüfung (Raudenbush/Bryk 2002). Darüber hinaus sollen Mehrebenenregressionsmodelle spezifiziert werden (Snijders/Bosker 2011), um die Bedingungsfaktoren eines effektiven bilingualen Geschichtsunterrichts auf Klassenebene identifizieren zu können.

Aspekte der Triangulation

Theoretische Vorüberlegungen

Der Begriff der Triangulation wurde in den letzten Jahren verstärkt in der empirischen Forschung sowie im Rahmen definitionstheoretischer Reflexionen verwandt (Gläser-Zikuda et al. 2012, Riemer 2011). Während Einigkeit darüber herrscht, dass Triangulation die Kombination verschiedener Zugänge zum Forschungsgegenstand meint, besteht Unbestimmtheit darüber, was genau in welchem Maße kombiniert wird.¹ Auch der Begriff „Triangulation“ ist nicht völlig eindeutig, sondern wird von verschiedenen Autoren in variierenden Bedeutungszusammenhängen und mitunter synonym mit dem „integrativen Ansatz“ nach Seipel/Rieker (2003) oder dem *mixed methods approach* nach Jick (1979) verwendet. Im Folgenden werden nun zwei theoretische Modelle der Triangulation vorgestellt, um diese dann in einem zweiten Schritt auf die COMBIH-Studie anzuwenden.

Ein Triangulationsmodell stammt von Ivankova/Creswell (2009). Ihr *mixed methods*-Ansatz schreibt nur solchen Forschungsprojekten triangulativen Charakter zu, die qualitativ und quantitativ gewonnene Datensätze miteinander kombinieren. Das *explanatory* und das *exploratory design* legen den Fokus auf nur eine Forschungsmethode (quantitativ oder qualitativ) und betrachten die jeweils andere lediglich als der Hauptstudie zuarbeitende Datenerhebungsmethode. Im *triangulation design* hingegen findet die qualitative und quantitative Datensammlung und -auswertung gleichberechtigt statt. In der Phase der Dateninterpretation werden die beiden Ansätze kombinativ zusammengeführt.

Das zweite Modell findet sich in Flick (2011), der an Denzins Unterteilung in vier Formen der Triangulation von 1970 anknüpft. Als triangulativ wird demnach angesehen,

- wenn mehrere Forscherinnen und Forscher eingesetzt werden, um Daten zu erheben (Forscher-Triangulation),
- wenn verschiedene Datensätze kombiniert werden, die an verschiedenen Orten, zu verschiedenen Zeitpunkten und an verschiedenen Personen(-gruppen) gewonnen wurden (Daten-Triangulation),

¹ Eine ausführliche Diskussion hierzu findet sich in Flick (2011).

- wenn sich Forscherinnen und Forscher mit mehreren, forschungsparadigmatisch divergierenden Perspektiven und Hypothesen einem Untersuchungsgegenstand nähern (Theorien-Triangulation)
- und wenn verschiedene methodische Zugänge miteinander kombiniert werden (Methoden-Triangulation), was sowohl innerhalb von als auch zwischen Forschungsparadigmen erfolgen kann (*within-method* bzw. *between-method*).

Die Kombination verschiedener Forschungszugänge verfolgt in der Regel zwei Absichten (Flick 2011: 9). Sie dient entweder der Validierung von Daten bzw. Erhebungsinstrumenten (hiermit soll Konvergenz erzielt werden) oder sie dient der Erkenntniserweiterung (dies führt entweder zu komplementären oder zu divergierenden Ergebnissen). Während erstere Funktion vor allem zu Beginn der Triangulationsforschung als Hauptziel gesehen wurde, fand mittlerweile ein Paradigmenwechsel statt, der nun die erweiterten Erkenntnismöglichkeiten in den Mittelpunkt rückt. Die vorliegende Studie folgt dem Konzept der Triangulation nach Flick, da sich Triangulation als übergreifender *umbrella term* verstehen lässt, innerhalb dessen sich Ivankovas/Creswells Ansatz zwar im Sinne der *between-method* verorten lässt, der darüber hinaus jedoch vielfältige Formen der Triangulation in Bezug auf Forscher, Datensätze und Theorien bietet.

Triangulation in der COMBIH-Studie

Forscher-Triangulation

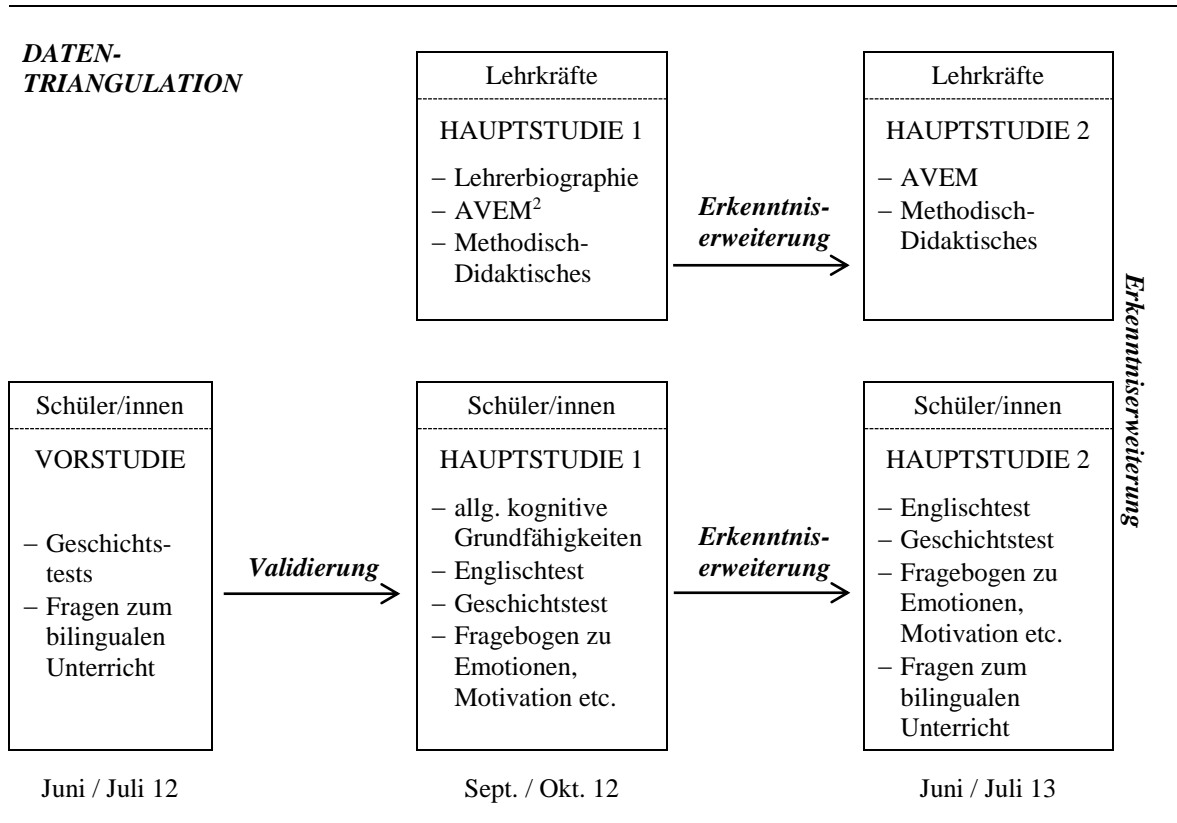
Um den bisherigen Forschungsstand zum bilingualen Sachfachunterricht systematisch zu erweitern, wurde die Expertise verschiedener Fachdisziplinen vereinigt. So wurden insbesondere die grundlegenden englisch- und geschichtsdidaktischen Fragestellungen der Studie sowie die Entwicklung der Leistungstests in den Sachfächern von den Fachdidaktikabteilungen Englisch und Geschichte der Pädagogischen Hochschule Ludwigsburg betreut. Für die zentralen psychologischen Konstrukte in den Fragebögen bedurfte es der Expertise der Pädagogischen Psychologie. Dies, sowie die Entwicklung des Studiendesigns, die Betreuung der psychometrischen Verfahren und statistisch-analytische Unterstützung wurde von der Abteilung Empirische Bildungsforschung und Pädagogische Psychologie der Universität Tübingen geleistet. Diese interdisziplinäre Zusammenarbeit im Rahmen des Kooperativen Promotionskollegs „Effektive Lehr-Lernarrangements“ führte zur Integration verschiedenster Perspektiven und Forschungstraditionen und somit zu einer möglichst angemessenen Annäherung an den Forschungsgegenstand.

Daten-Triangulation

Auch der Ansatz der Daten-Triangulation wurde in die COMBIH-Studie integriert. Dieser verfolgte zwei Absichten (vgl. Abb. 1). Erstens diente die Vorstudie der Validierung des neu entwickelten Test- und Befragungsmaterials. Nach der Auswertung der Pilotierungsdaten wurden die Aufgaben des Geschichtstests verbessert sowie eine Selektion der psychometrisch besten, den Gütekriterien empirischer Forschung entsprechenden Items zum bilingualen Unterricht vorgenommen. Diese so

optimierten Instrumente wurden dann in die Sammlung der übrigen (bereits in anderen Studien validierten) Instrumente für die Hauptstudie aufgenommen. Zweitens diente der Ansatz der Daten-Triangulation dazu, verschiedene Perspektiven miteinander zu kombinieren und somit eine Erkenntniserweiterung herbeizuführen. Dies wurde in COMBIH durch die folgenden fünf Maßnahmen erreicht.

Abb. 1: Daten-Triangulation in der COMBIH-Studie



Längsschnittliches Design

Schüler- und Lehrerdaten wurden zu zwei Messzeitpunkten erhoben. Der erste Messzeitpunkt dient der Erfassung der Ausgangslage der beiden Schülergruppen zu Beginn des bilingualen Unterrichts, die zweite Messung erlaubt Aussagen über mögliche Veränderungen. Mit dem Geschichtstest wird beispielsweise zunächst das Vorwissen der Schülerinnen und Schüler zu den Inhalten der Klasse 8 überprüft, um dann mithilfe der Messung am Ende des Schuljahres den Leistungszuwachs erfassen zu können. Dasselbe trifft auf die Entwicklung der Englischkompetenzen und der Schülermotivation und Emotionen bezüglich der Fächer Englisch und Geschichte zu. Würden die Schülerinnen und Schüler nur am Ende des Schuljahres, also nach erfolgtem bilingualen Geschichtsunterricht befragt und getestet, wüsste man zwar, wie die Lerner diese Unterrichtsform erlebt und welche Leistungen sie in Geschichte und Englisch erbracht haben. Es wären jedoch keine Informationen darüber

² Arbeitsbezogenes Verhaltens- und Erlebensmuster (vgl. Abschnitt 4.3)

vorhanden, ob diese Unterschiede nicht bereits vor Beginn des bilingualen Geschichtsunterrichts bestanden. Zwei Messzeitpunkte hingegen gestatten es – sofern die entscheidenden konfundierenden Variablen kontrolliert werden – Aussagen über die Effekte des bilingualen Geschichtsunterrichts zu treffen: Wie verändern sich Leistungen, Motivation und Emotionen von bilingual unterrichteten Lernern durch den bilingualen Geschichtsunterricht?

Leistungstests und Fragebogen

Die Schülerinnen und Schüler bearbeiteten zu beiden Messzeitpunkten Leistungstests und einen Fragebogen. Bezüglich der Englischleistungen wurde ein C-Test zur Erfassung der allgemeinen Englischkompetenz eingesetzt sowie ein Hörverstehenstest, da bezüglich dieses Kompetenzbereichs besonders große Unterschiede zwischen bilingual und einsprachig unterrichteten Schülerinnen und Schülern erwartet werden. Das Geschichtswissen wurde mithilfe eines Lückentexts und *Multiple Choice*-Aufgaben erhoben. Der Geschichtstest zu Inhalten aus Klasse 7 soll zeigen, wie gut die Lerner bisher in Geschichte waren. Der Test zu Inhalten aus Klasse 8 dient zum ersten Messzeitpunkt als Vorwissenstest und zu Messzeitpunkt zwei als Maß für den Zuwachs an Geschichtswissen. Um über das Sachwissen hinaus auch Kompetenzen im Fach Geschichte berücksichtigen zu können, wurden Fragen zur historischen De- und Rekonstruktionskompetenz integriert. Zusätzlich wurden als Indikatoren der Lernvoraussetzungen zwei kognitive Grundfähigkeitstests eingesetzt. Die Tests sind von besonderer Bedeutung, da fundierte, objektive Aussagen zu Schülerleistungen getroffen werden können, die mithilfe der Erhebung von Schulnoten und dem akademischen Selbstkonzept nicht möglich wären. Mit dem Schülerfragebogen wurden mehrere Aspekte erfasst: zum einen die zentralen Konstrukte Motivation und Emotionen in den Fächern Englisch und Geschichte. Zum anderen wurden Fragen integriert, die die Schüler dazu aufforderten, ihren Englisch- und Geschichtsunterricht sowie die Lehrkräfte einzuschätzen. So kann ein Bild davon gewonnen werden, wie Unterricht in den verschiedenen Klassen der Stichprobe ablief.³ Außerdem dienen diese Fragen der Kontrolle von Schülerleistung, Motivation und Emotionen. Mithilfe des umfangreichen Test- und Befragungsmaterials sollte der Forschungsgegenstand so umfassend und präzise wie möglich erfasst werden.

Schüler- und Lehrerperspektive

In der Studie wurden sowohl Schüler- als auch Lehrerdaten erhoben. Im Lehrerfragebogen befanden sich Fragen zu methodisch-didaktischen Aspekten des (bilingualen) Geschichtsunterrichts, Fragen zur Person, zur Lehrerbiographie, zu arbeitsbezogenen Verhaltens- und Erlebensmustern und eine Einschätzung der an der Studie teilnehmenden Klasse. Mithilfe dieser rund 160 Items sollte der Geschichtsunterricht der Schülerinnen und Schüler so ausgewogen und differenziert wie möglich

³ Einschränkung muss eingeräumt werden, dass lediglich die Wahrnehmung der Schüler (und Lehrkräfte) berichtet werden kann und unklar bleibt, inwieweit diese mit den tatsächlichen unterrichtlichen Vorgängen übereinstimmen.

erfasst werden. Weiterhin diene die Befragung der Lehrkräfte der Kontrolle von Kontextvariablen. Denn Lehrkräfte und die damit in Verbindung stehende Qualität des Unterrichts nehmen Einfluss sowohl auf motivationale und emotionale Aspekte als auch auf das Kompetenzerleben und die Leistungsentwicklung der Schülerinnen und Schüler (Kunter et al. 2011, Helmke/Schrader 2001). So können Aussagen über Motivation, Emotionen und Schülerkompetenzen getroffen werden, die unabhängig sind von mit der Lehrkraft in Verbindung stehenden Faktoren. Denn sollten diese Faktoren systematisch zwischen bilingual und einsprachig unterrichtenden Lehrkräften variieren, wäre dies eine weitere Varianzquelle, die es bei der Isolation der Effekte des bilingualen Unterrichts zu kontrollieren gilt.

Bilingual unterrichtete Klassen und zwei Vergleichsgruppen

Um die Effekte des bilingualen Geschichtsunterrichts auf die Entwicklung von Englisch- und Geschichtskompetenzen, Schülermotivation und Emotionen zu bestimmen, bedarf es mindestens einer Vergleichsgruppe deutschsprachig unterrichteter Schülerinnen und Schüler. Würden Ergebnisse auf eine deutlich positivere Entwicklung von Kompetenzen, Motivation und Emotionen bilingual unterrichteter Schülerinnen und Schüler in den Fächern Englisch und Geschichte verweisen, könnte dies das Konzept des bilingualen Lehrens und Lernens weiter stärken und die Ausweitung des bilingualen Unterrichtsangebots unterstützen. Durch die Einbindung von zwei Vergleichsgruppen – Parallelklassen und Klassen aus Schulen ohne bilinguales Angebot – wird bezweckt, mögliche Unterschiede innerhalb der nicht bilingual unterrichteten Schülerschaft zu untersuchen. Diese könnten zu Ungunsten der Parallelklassenschüler ausfallen und eine Konsequenz der positiven Selektion der bilingual unterrichteten Lerner an den bilingualen Zugschulen darstellen.

Gymnasien aus ganz Baden-Württemberg

Die an der Studie beteiligten Gymnasien sind über ganz Baden-Württemberg verteilt und tragen somit zum Ansatz der Daten-Triangulation bei, wonach verschiedene Datensätze an unterschiedlichen Orten gewonnen werden. Die Datenerhebung fand sowohl an Gymnasien statt, die an der hessischen Grenze liegen, als auch an Schulen am Bodensee und in Freiburg; an Gymnasien, die in der Nähe Frankreichs und in der Nähe Bayerns liegen. Darüber hinaus waren an der Studie sowohl städtische als auch ländliche Gymnasien beteiligt, durchschnittlich große Schulen mit rund 1000 Schülerinnen und Schülern und größere, bis zu elfzügige Schulen. Diese unterschiedlichen und recht gleichmäßig über Baden-Württemberg verstreut gelegenen Gymnasien decken ein weites Spektrum der Zielpopulation ab. Allerdings stellen die Schulen keine repräsentativ und zufällig gezogene Stichprobe dar (vgl. Abschnitt 4.1), weshalb die gewonnenen Daten nur begrenzt auf die Grundgesamtheit übertragen und verallgemeinert werden können.

Zusammenführung der triangulativen Aspekte

Bei der Analyse der Daten und Interpretation der gewonnenen Ergebnisse von COMBIH werden alle obig berichteten triangulativen Aspekte gleichermaßen berücksichtigt: die unterschiedlichen disziplinären Perspektiven, die beiden Messzeitpunkte, zu denen eine sehr große Stichprobe verteilt über das ganze Bundesland untersucht wurde, die objektiven Leistungstests sowie die standardisiert erfassten Selbstberichte zum Unterrichtserleben, zur Motivation und zu Emotionen, die Schüler- und Lehrerperspektive, sowie die unterschiedlichen Vergleichsgruppen. In diesem Sinne kann in Anlehnung an die Arbeiten von Ivankova/Creswell die Gewichtung der verschiedenen Datenquellen bei COMBIH – „the relative importance or priority given to each type of data“ (2009: 138) – als gleichgewichtet bezeichnet werden. Somit wird einer wichtigen Forderung dieses forschungsparadigmatischen Ansatzes Rechnung getragen, wonach der Mehrwert der Triangulation in der bedeutsamen Integration und nicht in einer bloßen Akkumulation von verschiedenen Zugängen zum Forschungsgegenstand liegt.

Literatur

- Abendroth-Timmer, Dagmar (2004). Evaluation bilingualer Module aus Schülerperspektive: zur Lernbewusstheit und ihrer motivationalen Wirkung. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 9/2, 1-27. Online: <http://zif.spz.tu-darmstadt.de/jg-09-2/beitrag/Abendroth2.htm> [23.10.2013].
- Alonso, Esmeralda / Grisaleña, Jesús / Campo, Alejandro (2008). Plurilingual Education in Secondary Schools: Analysis of Results. *International CLIL Research Journal* 1/1, 36-49. Online: <http://www.icrj.eu/11/article3.html> [27.08.2013].
- Badertscher, Hans / Bieri, Thomas (2009). *Wissenserwerb im Content and Language Integrated Learning. Empirische Befunde und Interpretationen*. Bern: Haupt.
- Baumert, Jürgen / Trautwein, Ulrich / Artelt, Cordula (2003). Schulumwelten. Institutionelle Bedingungen des Lehrens und Lernens. In: Baumert, Jürgen et al. (Hrsg.). *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen: Leske & Budrich, 259-330.
- Bertram, Christiane (2012). Wirksamkeit von Zeitzeugenbefragungen im Geschichtsunterricht – ein Beitrag für die empirische Geschichtsdidaktik. *Geschichte für heute. Zeitschrift für historisch-politische Bildung* 5/3, 21-33.
- Bonnet, Andreas (2012). Towards an Evidence Base for CLIL. How to Integrate Qualitative and Quantitative as well as Process, Product and Participant Perspectives in CLIL Research. *International CLIL Research Journal* 4/1, 66-78. Online: <http://www.icrj.eu/14/article7.html> [26.08.2013].
- Breidbach, Stephan / Viebrock, Britta (2012). CLIL in Germany: Results from Recent Research in a Contested Field of Education. *International CLIL Research Journal* 1/4, 5-16. Online: <http://www.icrj.eu/14/article1.html> [23.10.2013].
- Bruton, Anthony (2011). Is CLIL so Beneficial, or Just Selective? Re-evaluating Some of the Research. *System* 39/4, 523-532.
- Costa, Francesca / D'Angelo, Lauretta (2011). CLIL: A Suit for All Seasons. *Latin American Journal of Content & Language* 4/1, 1-13.
- Coyle, Do / Hood, Philip / Marsh, David C. (2010). *CLIL. Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Dallinger, Sara (2013). Die Einstellung der Eltern zum bilingualen Sachfachunterricht. Eine quantitative Studie an Realschulen in Baden-Württemberg. In: Hollm, Jan / Hüttermann, Armin / Keßler, Jörg-U. / Schlemminger, Gérald / Ade-Thurow, Benjamin (Hrsg.). *Bilinguales Lehren und Lernen in der Sekundarstufe I. Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik, 268-286.

- DESI-Konsortium (2006). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch Englisch Schülerleistungen International (DESI)*. Online: <http://www.dipf.de/de/projekte/pdf/biqua/desi-zentrale-befunde> [03.09.2013].
- Fehling, Sylvia (2008). *Language Awareness und bilingualer Unterricht. Eine komparative Studie*, Frankfurt/Main: Lang.
- Flick, Uwe (2011). *Triangulation. Eine Einführung*. 3. Aufl. Wiesbaden: VS Verlag.
- Gläser-Zikuda, Michaela / Seidel, Tina / Rohlf, Carsten / Gröschner, Alexander / Ziegelbauer, Sascha (Hrsg.) (2012). *Mixed Methods in der empirischen Bildungsforschung*. Münster: Waxmann.
- Götz, Thomas / Cronjäger, Hanna / Frenzel, Anne C. / Lüdtke, Oliver / Hall, Nathan C. (2010). Academic Self-concept and Emotion Relations: Domain Specificity and Age Effects. *Contemporary Educational Psychology* 35, 44-58.
- Götz, Thomas / Frenzel, Anne C. (2010). Über- und Unterforderungslangeweile im Mathematikunterricht. *Zeitschrift Empirische Pädagogik* 24/2, 113-134.
- Heller, Kurt A. / Perleth, Christoph (2000). *Kognitiver Fähigkeitstest für 4. – 12. Klassen. Revision (KFT 4-12+R)*. Göttingen: Hogrefe.
- Helmke, Andreas / Schrader, Friedrich-Wilhelm (2001). Determinanten der Schulleistung. In: Rost, Detlef H. (Hrsg.). *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz PVU, 81-90.
- Hollm, Jan / Hüttermann, Armin / Keßler, Jörg-U. / Schlemminger, Gérald (2010). *Zwischenbericht der wissenschaftlichen Begleitung zum Schulversuch „Bilinguale Züge an Realschulen“ in Baden-Württemberg. Projekttitle: BiliReal 2012*. Online: http://www.ph-ludwigsburg.de/fileadmin/subsites/2b-engl-t-01/Bili/Zwischenbericht_Schulversuch_BLLZuege_RS_Ba-Wue.pdf [10.09.2013].
- Hollm, Jan / Hüttermann, Armin / Keßler, Jörg-U. / Schlemminger, Gérald / Ade-Thurrow, Benjamin (Hrsg.) (2013). *Bilinguales Lehren und Lernen in der Sekundarstufe I. Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik.
- Ivankova, Nataliya V. / Creswell, John W. (2009). Mixed Methods. In: Heigham, Juanita / Crocker Robert A. (Hrsg.). *Qualitative Research in Applied Linguistics*. Basingstoke: Palgrave Macmillan, 135-161.
- Jick, Todd D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly* 24/4, 602-611.
- Kollenrott, Anne I. (2008). *Sichtweisen auf deutsch-englisch bilingualen Geschichtsunterricht. Eine empirische Studie mit Fokus auf interkulturelles Lernen*. Frankfurt/Main: Lang.
- Köller, Olaf / Leucht, Michael / Pant, Hans A. (2012). Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I. *Unterrichtswissenschaft* 40/4, 334-350.

- Krapp, Andreas (2002). Structural and Dynamic Aspects of Interest Development: Theoretical Considerations from an Ontogenetic Perspective. *Learning and Instruction* 12, 383-409.
- Kunter, Mareike / Baumert, Jürgen / Blum, Werner / Klusmann, Uta / Krauss, Stefan / Neubrand, Michael (Hrsg.) (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Lamsfuß-Schenk, Stefanie (2008). *Fremdverstehen im bilingualen Geschichtsunterricht. Eine Fallstudie*. Frankfurt/Main: Lang.
- Lasagabaster, David / Sierra, Juan M. (2009). Language Attitudes in CLIL and Traditional EFL Classes. *International CLIL Research Journal* 2/1, 4-17. Online: <http://www.icrj.eu/12/article1.html> [26.08.2013].
- Nold, Günter / Hartig, Johannes / Hinz, Silke / Rossa, Henning (2008). Klassen mit bilinguaem Sachfachunterricht: Englisch als Arbeitssprache. In: DESI-Konsortium (Hrsg.). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz, 451-457.
- Pekrun, Reinhard (2006). The Control-value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review* 18, 315-341.
- Pérez-Cañado, María L. (2012). CLIL Research in Europe: Past, Present, and Future. *International Journal of Bilingual Education and Bilingualism* 15/3, 315-341.
- Raudenbush, Stephen W. / Bryk, Anthony S. (2002). *Hierarchical Linear Models*. Thousand Oaks: Sage.
- Riemer, Claudia (2011). Empirie und Fremdsprachenforschung: Herausforderung Forschungsmethodik. In: Bausch, Karl-Richard / Burwitz-Melzer, Eva / Krumm, Hans-Jürgen (Hrsg.). *Fremdsprachen Lehren und Lernen: Rück- und Ausblick. Arbeitspapiere der 30. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts*. Tübingen: Narr, 194-203.
- Rumlich, Dominik (2013a). Students' General English Proficiency Prior to CLIL: Empirical Evidence for Substantial Differences Between Prospective CLIL and Non-CLIL Students in Germany. In: Breidbach, Stephan / Viebrock, Britta (Hrsg.). *Content and Language Integrated Learning (CLIL) in Europe. Research Perspectives on Policy and Practice*, Frankfurt/Main: Lang, 181-201.
- Rumlich, Dominik (2013b). *The Development of CLIL and Non-CLIL Students' General Language Proficiency: Findings from the Large-scale Longitudinal Study DENOCS*. Vortrag auf dem Kongress der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF) in Augsburg am 27.09.2013.
- Rumlich, Dominik (2014). Prospective CLIL and Non-CLIL Students' Interest in English (Classes): A Quasi-experimental Study on German Sixth-graders. In: Breeze, Ruth / Martínez Pasamar,

- Concepción / Llamas Saíz, Carmen / Tabernero Sala, Cristina (Hrsg.). *Integration of Theory and Practice in CLIL*. Amsterdam: Rodopi.
- Schaarschmidt, Uwe / Fischer, Andreas W. (2008). *Arbeitsbezogenes Verhaltens- und Erlebensmuster (AVEM)*. Frankfurt/Main: Pearson.
- Schreiber, Waltraud / Körber, Andreas / Borries van, Bodo / Krammer, Reinhard / Leutner-Ramme, Sibylla / Mebus, Sylvia / Schöner, Alexander / Ziegler, Béatrice (2006). *Historisches Denken. Ein Kompetenz-Strukturmodell*. Neuried: ars una.
- Seipel, Christian / Rieker, Peter (2003). *Integrative Sozialforschung. Konzepte und Methoden der qualitativen und quantitativen empirischen Forschung*. Weinheim: Juventa.
- Snijders, Tom / Bosker, Roel (2011). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. Los Angeles: Sage.
- Wigfield, Allan L. / Eccles, Jacquelynne S. (2000). Expectancy-value Theory of Achievement Motivation. *Contemporary Educational Psychology* 25, 68-81.
- Wild, Elke / Hofer, Manfred (2000). Elterliche Erziehung und die Veränderung motivationaler Orientierungen in der gymnasialen Oberstufe und der Berufsschule. In: Schiefele, Ulrich / Wild, Klaus-Peter (Hrsg.). *Interesse und Lernmotivation*. Münster: Waxmann, 31-52.
- Witzigmann, Stéfanie (2011). *Bildende Kunst in der Zielsprache Französisch als Einstieg ins bilinguale Lehren und Lernen. Eine explorative Studie in einer fünften Realschulklasse*. Frankfurt/Main: Lang.
- Zydati, Wolfgang (2007). *Deutsch-Englische Zge in Berlin (DEZIBEL). Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien*. Frankfurt/Main: Lang.
- Zydati, Wolfgang (2012). Linguistic Thresholds in the CLIL Classroom? The Threshold Hypothesis Revisited. *International CLIL Research Journal* 4/1, 17-28. Online: <http://www.icrj.eu/14/article2.html> [26.08.2013].

Lernausgangslagen von Schülerinnen und Schülern zu Beginn des bilingualen Geschichtsunterrichts

Sara Dallinger

Abteilung Englisch, Pädagogische Hochschule Ludwigsburg

Kathrin Jonkmann

Abteilung Empirische Bildungsforschung und Pädagogische Psychologie, Universität Tübingen
Lehrgebiet Bildungspsychologie, FernUniversität Hagen

Dallinger, Sara & Jonkmann, Kathrin (2014). Lernausgangslagen von Schülerinnen und Schülern zu Beginn des bilingualen Geschichtsunterrichts. In Christiane Fäcke, Martina Rost-Roth & Engelbert Thaler (Hrsg.), *Sprachenausbildung – Sprachen bilden aus – Bildung aus Sprachen: Dokumentation zum 25. Kongress für Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF) Augsburg, 25.-28. September 2013*. Baltmannsweiler: Schneider Verlag Hohengehren, 173–183. (Die Zitation folgt den Vorgaben des Sammelbands.)

Lernausgangslagen von Schülerinnen und Schülern zu Beginn des bilingualen Geschichtsunterrichts

Einleitung

Der bilinguale Sachfachunterricht steht bereits seit über einem Jahrzehnt im Mittelpunkt regen Forschungsinteresses (Pérez-Cañado 2012). Viele Studien kommen zu überwiegend positiven Ergebnissen: Schülerinnen und Schüler scheinen vom bilingualen Sachfachunterricht zu profitieren (Nold, Hartig, Hinz & Rossa 2008; Zydatið 2007) und sowohl Lehrkräfte (Kollenrott 2008), Schulverwaltung (Hollm, Hüttermann, Keßler, Schlemminger & Ade-Thurow 2013) als auch Eltern (Dallinger 2013) bewerten diese Unterrichtsform sehr positiv. In jüngerer Zeit mehrten sich jedoch auch kritische Einschätzungen (Bonnet 2012; Pérez-Cañado 2012). Bruton (2011) gab zu bedenken, dass die vielfach nachgewiesenen positiven Leistungsergebnisse bilingual unterrichteter Lerner dadurch bedingt sein könnten, dass bilinguale Angebote vorwiegend von einer besonders motivierten und leistungsstarken Schülerklientel aus bildungsnahen Elternhäusern wahrgenommen werden. Leistungsunterschiede zwischen bilingual und einsprachig unterrichteten Lernern wären demnach weitestgehend auf deren unterschiedliche Lernausgangslagen und nicht auf die unterschiedliche Beschulung zurückzuführen. Bonnet (2012: 67) argumentierte ähnlich: "There are strong indications from evaluation research [...] that social selection plays an important role in German CLIL. Therefore, high outcomes in CLIL groups might well be ascribed to factors other than classroom practice"⁴. Zur Abschätzung der Effektivität des bilingualen Unterrichts ist es also zunächst erforderlich, die potentiell unterschiedlichen Lernausgangslagen bilingual und regulär einsprachig unterrichteter Schülerinnen und Schüler festzustellen. Dies ist das Anliegen des folgenden Beitrags.

Auswahlverfahren für bilinguale Züge

Die von Bonnet (2012: 67) erwähnte "social selection" wird durch die Auswahlverfahren der Schulen, die über ein fest eingerichtetes bilinguales Profil verfügen, begünstigt. Schülerinnen und Schüler, die im Anschluss an die Grundschule an einem solchen bilingualen Zug teilnehmen wollen, werden meist anhand der Grundschulnoten und einem Gespräch zwischen Eltern und Schulleitung der bilingual unterrichtenden Schule ausgewählt (Zydatið 2007: 70). Hierbei wird den Eltern häufig vermittelt, dass der bilinguale Sachfachunterricht insbesondere für sehr motivierte Kinder mit hoher Lernbereitschaft, die höheren kognitiven Ansprüchen und einem gesteigerten Lernaufwand gewachsen sind, geeignet sei (Rumlich 2013: 184). Diese Mechanismen der Selbst- und Fremdselektion sollten dazu führen, dass bilingual unterrichtete Schülerinnen und Schüler besonders günstige Lernvoraussetzungen aufweisen hinsichtlich der Lernmotivation (Wigfield & Eccles 2000), der kognitiven Grundfähigkeiten (Helmke & Schrader 2001: 81) und der Bildungsnähe des

⁴ CLIL = *Content and Language Integrated Learning*

Elternhauses (Köller, Leucht & Pant 2012: 339). Darüber hinaus müssen die unterschiedlich ausgebildeten Englischkompetenzen und das Vorwissen im Sachfach vor Beginn des bilingualen Unterrichts berücksichtigt werden, um Aussagen über den inkrementellen Leistungszuwachs nach erfolgtem (bilinguaem) Unterricht treffen zu können. Im Folgenden wird eine Auswahl an Studien präsentiert, mithilfe derer Rückschlüsse auf die Lernausgangslagen bilingual und einsprachig unterrichteter Lerner gezogen werden können.

Lernausgangslagen bilingual unterrichteter Lerner

Rumlich (2013) untersuchte in DENOCS die Englischkompetenz (C-Test) von 340 bilingual und 430 einsprachig unterrichteten Gymnasiastinnen und Gymnasiasten. Am Ende der Klasse 6, nach verstärktem Englischunterricht in den Klassenstufen 5 und 6, vor Beginn des bilingualen Sachfachunterrichts, zeigten die zukünftig bilingual unterrichteten Lerner deutlich bessere Englischleistungen sowie ein höheres selbstberichtetes Interesse an der Fremdsprache und am Fach Englisch als die einsprachig unterrichtete Schülergruppe (Rumlich 2014). Fehling (2008) untersuchte die *language awareness* und die verbalen kognitiven Grundfähigkeiten von 74 bilingual und 65 einsprachig unterrichteten Gymnasiastinnen und Gymnasiasten zu Beginn des bilingualen Unterrichts Anfang Klasse 7. In drei von vier gemessenen Dimensionen der *language awareness* (affektiv, kognitiv, politisch) sowie in den verbalen Grundfähigkeiten erzielte die bilingual unterrichtete Schülergruppe bessere Resultate.

Bisher sind keine Studien bekannt, die das Vorwissen in einem Sachfach erhoben haben, bevor dieses bilingual unterrichtet wird. Die wenigen vorliegenden Befunde (Bonnet 2012: 67) beziehen sich auf querschnittliche Studien mit Datenerhebungen nach erfolgtem bilinguaem Unterricht. Madrid (2011) und Passon (2007) erhoben die Geographiekompetenzen von Grund- und Sekundarstufenschülerinnen und -schülern und kamen zu dem Ergebnis, dass die einsprachig und bilingual unterrichteten Lerner vergleichbare Leistungen erbrachten. Auch Alonso, Grisaleña & Campo (2008) resümierten, dass die bilingual unterrichteten Lerner in den Sachfächern ähnliche Leistungen wie die Vergleichsgruppe zeigten. Diese Aussage stützte sich jedoch lediglich auf die subjektive Einschätzung von Lehrkräften sowie auf die Noten der Lerner, die verzerrenden Referenzgruppeneffekten unterliegen können (Baumert et al. 2003). Ähnlich dünn ist die empirische Basis zur Motivation im Sachfach (Lasagabaster & Sierra 2009: 4). Verschiedene Studien kamen mithilfe einmaliger Datenerhebungen zu dem Schluss, dass sich bilingualer Sachfachunterricht positiv auf die Motivation der Lerner auswirke (z.B. Abendroth-Timmer 2007; Lamsfuß-Schenk 2008). Da jedoch die Ausgangslagen der beiden Schülergruppen nicht untersucht wurden, lassen sich mögliche Eingangsunterschiede nicht von Effekten des bilingualen Unterrichts separieren.

Köller et al. (2012) berücksichtigten den sozioökonomischen Status (SES) als wichtigen Aspekt der Lernvoraussetzungen. Anhand der Angaben zu den Elternberufen und den Bildungsabschlüssen wurden der *International Socio-Economic Index of Occupational Status* (ISEI;

Ganzeboom, de Graaf, Treiman & de Leeuw 1992) sowie das Bildungsniveau (ISCED-97; UNESCO 1997) von 345 bilingual und 9521 einsprachig unterrichteten Lernern der 9. Klassenstufe bestimmt. Es zeigten sich keine statistisch bedeutsamen Unterschiede zwischen den Schülergruppen (vgl. Zydati 2007). Allerdings wurden bilingual unterrichtete Schülerinnen und Schüler mit Lernern aus Schulen ohne bilinguales Angebot verglichen. Betrachtet man jedoch die Verhältnisse *innerhalb* von Schulen mit bilingualem Zug, ist zu erwarten, dass es zu Selektionseffekten hinsichtlich des familiären Hintergrundes kommt.

Es lässt sich zusammenfassen, dass bedeutsame Eingangsunterschiede zugunsten der bilingual unterrichteten Schülerinnen und Schüler hinsichtlich der Englischkompetenzen und -motivation sowie der kognitiven Grundfähigkeiten bestehen. Für das Sachfach liegen bisher keine Befunde zu den Ausgangsleistungen bilingual und einsprachig Unterrichteter vor. Hinsichtlich des SES wurden in bisherigen Untersuchungen keine Unterschiede zwischen bilingual und einsprachig unterrichteten Lernern gefunden.

Forschungsfragen

Angesichts dieser Befundlage bedarf es weiterer Forschung, die die Lernausgangslagen von bilingual und einsprachig unterrichteten Schülerinnen und Schülern mittels standardisierter Leistungstests und Fragebögen bestimmt und zwar sowohl bezüglich der Leistungen und Motivation in Englisch und im Sachfach als auch hinsichtlich wichtiger Lernprädiktoren wie den verbalen und figuralen kognitiven Grundfähigkeiten und dem sozioökonomischen Status (SES). Dies stellt einen Forschungsschwerpunkt der Studie COMBIH ("Competences and Motivation in Bilingual Instruction in History") dar. Zur Untersuchung der Ausgangslagen wurde der erste Messzeitpunkt der Längsschnittstudie zu Beginn des bilingualen Geschichtsunterrichts in Klassenstufe 8 betrachtet. COMBIH vergleicht bilingual unterrichtete Lerner sowohl mit ihren Parallelklassen als auch mit Schülerinnen und Schülern aus Schulen ohne bilingualen Zug. In Bezug auf die einsprachig unterrichteten Parallelklassen von Schulen mit bilingualem Zug wird erwartet, dass die bilingual unterrichteten Lerner über einen höheren SES verfügen, während sich im Vergleich zur zweiten einsprachig unterrichteten Gruppe deutlich geringere Unterschiede zeigen sollten. Hinsichtlich der kognitiven Grundfähigkeiten werden insbesondere für die verbalen Fähigkeiten hohe Leistungen bei den bilingual unterrichteten Schülerinnen und Schülern erwartet. Für das Fach Englisch werden für sie ebenfalls bessere Leistungen und eine höhere Motivation erwartet. Aufgrund der besseren Lernvoraussetzungen sowie der Zuordnung des Fähigkeitsselbstkonzepts Geschichte zum verbalen Selbstkonzept (Marsh 1990) – ein Bereich, für den Vorteile zugunsten der Bilingual-Klassen erwartet werden – sollten sich auch im Fach Geschichte Leistungs- und Motivationsvorsprünge zugunsten der Bilingual-Klassen zeigen.

Methode

Die Stichprobe der Studie umfasst 1806 Schülerinnen und Schüler der 8. Klassenstufe aus 75 Klassen in 37 Gymnasien in Baden-Württemberg (50.3% Mädchen; $M_{\text{Alter}}=13.5$, $SD_{\text{Alter}}=.51$). Die Stichprobe gliedert sich in drei Teilstichproben auf. Die erste Teilstichprobe bilden 703 Schülerinnen und Schüler aus 30 Klassen (28 Gymnasien), die ab Klasse 5 den bilingualen Zug ihrer Schule besuchten, in den Klassenstufen 5 und 6 am verstärkten Englischunterricht von zwei zusätzlichen Wochenstunden teilnahmen und ab Klasse 7 in Geographie bilingual unterrichtet wurden. Die 659 Schülerinnen und Schüler der Vergleichsgruppe I (28 Klassen) besuchten ihre Parallelklassen ohne bilingualen Unterricht. Vergleichsgruppe II umfasste 444 Schülerinnen und Schüler aus 17 Klassen (9 Gymnasien) ohne bilinguales Unterrichtsangebot. Zur Stichprobenrekrutierung wurden alle 65 Gymnasien mit bilingualem Zug in Baden-Württemberg gezielt kontaktiert und um Teilnahme gebeten. Bei den Schulen der Vergleichsgruppe II handelt es sich aus Machbarkeitsgründen um eine Gelegenheitsstichprobe aus dem näheren Umkreis der Hochschulen.

Der erste Messzeitpunkt von COMBIH fand im Herbst 2012 zu Beginn der 8. Klassenstufe statt. Vor Beginn der Datenerhebungen wurden alle Materialien und Prozeduren vom Ministerium für Kultus, Jugend und Sport Baden-Württemberg für die Untersuchung freigegeben. Alle teilnehmenden Schülerinnen und Schüler mussten über eine Elterneinverständniserklärung verfügen. Die Datenerhebung belief sich auf 135 Minuten plus Pausenzeiten und wurde mithilfe eines standardisierten Manuals von geschulten Testleitern durchgeführt.

Die Untersuchungsmaterialien der Schülerinnen und Schüler umfassten Leistungstests und einen Fragebogen. Die Englischleistungen wurden mittels eines eigens entwickelten C-Tests (70 Items, $\alpha = .89$) sowie mit einem Hörverstehenstest aus dem Aufgabenpool der Bildungsstandards (IQB; Köller, Knigge & Tesch 2010) geprüft (27 Items, $\alpha = .73$). Das Geschichtswissen aus Klasse 7 wurde mithilfe eines eigens entwickelten Lückentexts und Multiple Choice-Aufgaben gemessen (25 Items, $\alpha = .72$). Die figuralen und verbalen kognitiven Grundfähigkeiten wurden mit Subtests des Kognitiven Fähigkeitstests erhoben (KFT; Heller & Perleth 2000; KFT figural: 25 Items, $\alpha = .91$; KFT verbal: 20 Items, $\alpha = .72$). Die Messung der Schülermotivation in den Fächern Englisch und Geschichte orientierte sich theoretisch am Erwartungs-Mal-Wert-Modell der Leistungsmotivation von Eccles (z.B. Wigfield & Eccles 2000). Es umfasst vier Wertkomponenten und das Fähigkeitsselbstkonzept: *Interesse* (z.B. "Englisch ist eines meiner Lieblingsfächer", 5 Items, ENG: $\alpha = .72$, GES: $\alpha = .82$), *Wichtigkeit* (z.B. "Ich halte das Fach Geschichte für sehr wichtig", 3 Items, ENG: $\alpha = .83$, GES: $\alpha = .80$), *Nützlichkeit* (z.B. "Ich werde in meinem Leben Englisch noch oft benötigen", 5 Items, ENG: $\alpha = .88$, GES: $\alpha = .86$), *Kosten* (z.B. "In Geschichte muss ich viel Zeit aufwenden, um eine gute Note zu bekommen", 2 Items, ENG: $\alpha = .82$, GES: $\alpha = .77$), *Fähigkeitsselbstkonzept* (z.B. "Ich bin gut in Englisch", 4 Items, ENG: $\alpha = .84$, GES: $\alpha = .86$). Alle Items zur Motivation waren auf einer vierstufigen Likert-Skala zu beantworten (1 = "stimmt gar nicht", 4 = "stimmt genau"). Die Operationalisierung des SES orientierte sich an

Schulleistungsstudien wie PISA. Mithilfe der Angaben zum elterlichen Schul- und Bildungsabschluss und dem Beruf wurde der *International Socio-Economic Index of Occupational Status* (ISEI; Ganzeboom et al. 1992; HISEI für den höheren ISEI der beiden Elternteile) gebildet. Dieser bewegte sich zwischen 11 (sehr niedriger SES) und 88 (sehr hoher SES).

Die inferenzstatistischen Analysen (Mittelwertvergleiche) wurden in Mplus 7.1 (Muthén & Muthén 1998-2012) durchgeführt. Durch die Wahl der Option *type = complex* wurde die Mehrebenenstruktur der Stichprobe (Nestung der Schüler in ihren Klassen) berücksichtigt, die dazu führt, dass keine vollständig unabhängigen Beobachtungen vorliegen. Unterbliebe dies, würden die Standardfehler unterschätzt und es erfolgte eine zu liberale statistische Signifikanzprüfung (Raudenbush & Bryk 2002). Einzelne fehlende Werte im Datensatz wurden in Mplus mithilfe des *Full Information Maximum Likelihood*-Verfahrens (FIML) geschätzt.

Ergebnisse

Tabelle 1 berichtet die Mittelwerte und Standardabweichungen der abhängigen Variablen sowie die Ergebnisse der Signifikanztests der Gruppenvergleiche. Die Mittelwertunterschiede zwischen den beiden einsprachig unterrichteten Vergleichsgruppen erwiesen sich in keinem Fall als statistisch bedeutsam und werden daher nicht berichtet.

Wie erwartet verweist Tabelle 1 auf einen höheren SES der bilingual unterrichteten Schülerinnen und Schüler gegenüber den einsprachig unterrichteten Lernern aus den Parallelklassen. Der Mittelwertunterschied beläuft sich auf eine fünftel Standardabweichung ($d = .20$), was nach Cohen (1988: 24-26) einem schwachen Effekt entspricht. Zur zweiten Vergleichsgruppe, den Lernern aus Schulen ohne bilinguales Angebot, finden sich erwartungsgemäß keine statistisch bedeutsamen Unterschiede. Weiterhin bestehen erwartungskonform signifikante Unterschiede zugunsten der Bilingual-Klassen bezüglich der verbalen kognitiven Grundfähigkeiten. Der Mittelwertunterschied zu den Parallelklassen beträgt rund eine drittel Standardabweichung ($d = .36$) und zu den Lernern aus Schulen ohne bilinguales Angebot knapp eine viertel Standardabweichung ($d = .24$). Somit ergeben sich nach Cohen zwar nur kleine Effekte, die jedoch als praktisch sehr bedeutsam angesehen werden dürfen. Bezüglich des figuralen kognitiven Grundfähigkeitstests zeigten sich erwartungsgemäß keine statistisch bedeutsamen Gruppenunterschiede. In Englisch schnitten die bilingual unterrichteten Lerner wie erwartet deutlich besser ab als die Vergleichsgruppen. Der Mittelwertunterschied hinsichtlich des C-Tests beläuft sich für die Parallelklassen auf knapp eine Standardabweichung ($d = .94$) und für die zweite Vergleichsgruppe auf eine dreiviertel Standardabweichung ($d = .76$). Bezüglich des Hörverstehenstests waren die Bilingual-Klassen um gut zwei drittel Standardabweichungen besser als die Parallelklassen ($d = .68$) und um eine dreiviertel Standardabweichung ($d = .76$) besser als die Lerner aus Schulen ohne bilinguales Angebot. Die Autoren des Ländervergleichs von 2009 (Köller et al. 2010) berichten, dass der Leistungszuwachs im Fach Englisch in der Sekundarstufe I in einem Schuljahr einem Effekt von ungefähr 0.40 Standardabweichungen entspricht. Damit beträgt der vorliegende Leistungsvorsprung

der bilingual unterrichteten Lerner im Fach Englisch ungefähr eineinhalb bis zwei Schuljahre. Auch hinsichtlich des Fachs Geschichte zeigten die Bilingual-Klassen im Vorwissenstest eine um knapp eine drittel Standardabweichung ($d = .29$) bessere Leistung als die einsprachig unterrichteten Parallelklassen, was auf einen Leistungsvorsprung von etwa einem dreiviertel Schuljahr hindeutet. Hinsichtlich der zweiten Vergleichsgruppe ergaben sich keine statistisch signifikanten Unterschiede. Die bilingual unterrichteten Lerner weisen des Weiteren in Englisch in allen fünf Motivationskalen positivere Werte auf als die einsprachig unterrichteten Gruppen. Die Unterschiede liegen in einem Bereich zwischen einer viertel und einer halben Standardabweichung ($.23 < d < .57$). Für Geschichte zeigten sich die bilingual unterrichteten Lerner in den Subkategorien *Wichtigkeit*, *Nützlichkeit* und *Selbstkonzept* um rund eine viertel Standardabweichung ($.13 < d < .26$) motivierter als die Lerner der einsprachig unterrichteten Schülergruppen. Es ergeben sich also kleine bis mittelgroße Effekte zugunsten der Bilingual-Klassen.

Tab.1: Unterschiede in den Lernausgangslagen zw. bilingual und einsprachig unterrichteten Lernern

| | BILI | | Mono1 | | Mono2 | | Vergleich BILI vs. | |
|-------------------|----------|-----------|----------|-----------|----------|-----------|-----------------------|-------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | Mono1 | Mono2 |
| HISEI | 65.43 | 17.73 | 61.87 | 18.70 | 64.46 | 17.06 | ** | ns |
| KFT verbal | 13.88 | 3.39 | 12.62 | 3.46 | 13.04 | 3.47 | *** | *** |
| KFT figural | 17.51 | 6.25 | 16.55 | 6.43 | 16.80 | 6.26 | ns | ns |
| <i>Englisch</i> | | | | | | | | |
| C-Test | 44.36 | 8.55 | 35.61 | 8.52 | 37.28 | 8.72 | *** | *** |
| Hörverstehen | 22.53 | 2.53 | 20.31 | 3.39 | 20.09 | 3.30 | *** | *** |
| Interesse | 2.98 | 0.62 | 2.72 | 0.67 | 2.80 | 0.66 | *** | ** |
| Wichtigkeit | 3.53 | 0.54 | 3.30 | 0.63 | 3.36 | 0.64 | *** | *** |
| Nützlichkeit | 3.65 | 0.48 | 3.45 | 0.61 | 3.52 | 0.59 | *** | ** |
| Kosten | 2.89 | 0.79 | 2.57 | 0.78 | 2.57 | 0.83 | *** | *** |
| Selbstkonzept | 3.44 | 0.57 | 3.06 | 0.68 | 3.12 | 0.70 | *** | *** |
| <i>Geschichte</i> | | | | | | | | |
| Wissen Kl. 7 | 11.40 | 3.94 | 10.20 | 4.34 | 10.96 | 3.90 | * | ns |
| Interesse | 2.26 | 0.69 | 2.16 | 0.76 | 2.16 | 0.73 | ns | ns |
| Wichtigkeit | 2.77 | 0.66 | 2.59 | 0.69 | 2.64 | 0.69 | *** | * |
| Nützlichkeit | 2.47 | 0.64 | 2.29 | 0.69 | 2.34 | 0.70 | ** | * |
| Kosten | 2.71 | 0.78 | 2.68 | 0.80 | 2.61 | 0.80 | ns | ns |
| Selbstkonzept | 3.00 | 0.66 | 2.84 | 0.73 | 2.81 | 0.77 | ** | *** |

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, BILI = Bilingual-Klassen, Mono1 = einsprachig unterrichtete Parallelklassen, Mono2 = Klassen aus Schulen ohne bilingualen Zug, KFT = Kognitiver Grundfähigkeitstest, HISEI = *International Socio-Economic Index of Occupational Status*.
 * $p < .05$, ** $p < .01$, *** $p < .001$.

Zusammenfassung und Diskussion

Die vorliegende Studie untersuchte Unterschiede in den Lernausgangslagen bilingual und einsprachig unterrichteter Schülerinnen und Schüler zu Beginn des bilingualen Geschichtsunterrichts

in Klasse 8. Es zeigte sich, dass bilingual unterrichtete Lerner über bessere verbale kognitive Grundfähigkeiten verfügen als die Vergleichsgruppen sowie über einen höheren SES als die einsprachig unterrichteten Parallelklassen. Kleine Unterschiede hinsichtlich des figuralen Tests waren statistisch nicht signifikant. Bezüglich der Englischtests weisen die bilingual unterrichteten Schülerinnen und Schüler große Leistungsvorsprünge von bis zu zwei Schuljahren auf. Auch im Geschichtstest zeigten die bilingual unterrichteten Lerner bessere Leistungen im Umfang eines dreiviertel Schuljahres gegenüber den einsprachig unterrichteten Parallelklassen. Weiterhin zeigt sich die bilingual unterrichtete Gruppe in Englisch und in etwas geringerem Maße auch in Geschichte motivierter als die einsprachig unterrichteten Lerner. Zwischen den beiden Vergleichsgruppen ergaben sich keine statistisch bedeutsamen Unterschiede.

Köller et al. (2012) zeigten keine Unterschiede im SES zwischen bilingual und einsprachig unterrichteten Schülergruppen auf. Dem zugrunde liegt ein Vergleich zwischen Schülerinnen und Schülern von Schulen mit und ohne bilingualen Zug. Dieser Befund wird von COMBIH bestätigt. Darüber hinaus kann die vorliegende Studie jedoch auch zeigen, dass es innerhalb eines Gymnasiums mit bilingualem Zug sehr wohl positive Selektionseffekte zugunsten der bilingual unterrichteten Lerner gibt, dass diese also über einen höheren SES verfügen als ihre einsprachig unterrichteten Parallelklassen. Bezüglich der kognitiven Grundfähigkeiten konnte COMBIH Fehlings (2008) Befunde replizieren, wonach Bilingual-Klassen im verbalen Bereich besser abschnitten als einsprachig unterrichtete Lerner. Darüber hinaus wird gezeigt, dass keine statistisch bedeutsamen Unterschiede im Bereich der figuralen kognitiven Grundfähigkeiten bestehen.

Der große Leistungsvorsprung der Bilingual-Klassen in den Englischtests ist ein bis dato vielfach repliziertes Ergebnis (Köller et al. 2012; Nold et al. 2008; Rumlich 2013). Er lässt sich vermutlich größtenteils durch den verstärkten Englischunterricht für Schülerinnen und Schüler eines bilingualen Zugs in den Klassenstufen 5 und 6 erklären (vgl. Bos, Bonsen & Gröhlich 2009). Die guten Englischleistungen könnten weiterhin auch durch deren hohe Motivation für das Fach Englisch bedingt sein. Darauf verweisen insbesondere die starken Effekte in Bezug auf das fachspezifische Selbstkonzept, dessen reziproker Zusammenhang mit Schülerleistung bereits in vielen Studien nachgewiesen wurde (Valentine, DuBois & Cooper 2004: 126f.). Eine höhere Motivation von bilingual unterrichteten Lernern im Fach Englisch (vgl. Rumlich 2014) überrascht nicht, da sich diese bewusst für ein bilinguales Profil und somit für ein Mehr an Englisch entschieden haben. Im Rahmen längsschnittlicher Studien müsste untersucht werden, ob vom bilingualen Unterricht weitere motivationsförderliche Wirkungen ausgehen.

Der Leistungsvorsprung der bilingual unterrichteten Schülerinnen und Schüler gegenüber ihren einsprachig unterrichteten Parallelklassen im Geschichtstest geht vermutlich auf die besseren Lernvoraussetzungen zurück. Wie bereits gezeigt, ergibt sich innerhalb eines Gymnasiums mit bilingualem Zug ein höherer SES für die Teilnehmer dieses Zugs. Außerdem verfügen die bilingual unterrichteten Lerner über bessere verbale kognitive Grundfähigkeiten als die Vergleichsgruppen.

Beides sind Aspekte, die in einem positiven Zusammenhang mit Schülerleistung stehen (Helmke & Schrader 2001). Dass sich keine Unterschiede in den Geschichtsleistungen im Vergleich zu den Schülerinnen und Schülern von Gymnasien ohne bilinguales Angebot zeigen, könnte darauf zurückgehen, dass sich deren Lernvoraussetzungen bezüglich des SES nicht so sehr von denen der bilingual unterrichteten Schülergruppe unterscheiden. Die höhere Motivation der Bilingual-Klassen in Geschichte ist insofern zu erwarten, als dass dieses Fach einen starken verbalen Anteil aufweist. Dies zeigt sich insbesondere in den Unterschieden im *Fähigkeitsselbstkonzept* in Geschichte, das zum verbalen Selbstkonzept gehört (Marsh 1990).

Zusammenfassend verweisen die vorliegenden Befunde auf substantielle Unterschiede zugunsten der bilingual unterrichteten Schülerinnen und Schüler zu Beginn des bilingualen Geschichtsunterrichts in Klasse 8. Dies trifft sowohl auf den (fremd-)sprachlichen Bereich als auch auf andere verbale Domänen (Fach Geschichte) zu. Diese positivere Ausgangslage bilingual unterrichteter Lerner scheint durch grundsätzlich günstigere Lernvoraussetzungen bedingt zu sein sowie durch eine personenimmanente und durch vermehrten Englischunterricht herbeigeführte, gesteigerte Sprachaffinität und Sprachkompetenz. Diese Befunde stellen die Basis für den Längsschnitt von COMBIH dar, der diese Ausgangslagen berücksichtigt, um Aussagen über Veränderungseffekte bilingualen Unterrichts treffen zu können.

Literaturverzeichnis

- Abendroth-Timmer, Dagmar (2007), *Akzeptanz und Motivation. Empirische Ansätze zur Erforschung des unterrichtlichen Einsatzes von bilingualen und mehrsprachigen Modulen*. Frankfurt/ Main: Peter Lang.
- Alonso, Esmeralda; Grisaleña, Jesús & Campo, Alejandro (2008), Plurilingual education in secondary schools: Analysis of results. *International CLIL Research Journal* 1: 1, 36-49. [<http://www.icrj.eu/11/article3.html>].
- Baumert, Jürgen; Trautwein, Ulrich & Artelt, Cordula (2003), Schulumwelten. Institutionelle Bedingungen des Lehrens und Lernens. In: Baumert, Jürgen; Artelt, Cordula; Klieme, Eckhard et al. (Hrsg.) (2003), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen: Leske/Budrich, 259-330.
- Bonnet, Andreas (2012), Towards an evidence base for CLIL. How to integrate qualitative and quantitative as well as process, product and participant perspectives in CLIL research. *International CLIL Research Journal* 4: 1, 66-78. [<http://www.icrj.eu/14/article7.html>].
- Bos, Wilfried; Bosen, Martin; Gröhlich, Carola (2009), *KESS 7. Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*. Münster: Waxmann.
- Bruton, Anthony (2011), Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System* 39: 4, 523-532.
- Cohen, Jacob (1988), *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- Dallinger, Sara (2013), Die Einstellung der Eltern zum bilingualen Sachfachunterricht. Eine quantitative Studie an Realschulen in Baden-Württemberg. In: Hollm, Jan et al. (Hrsg.) (2013), *Bilinguales Lehren und Lernen in der Sekundarstufe I. Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik, 268-286.
- Fehling, Sylvia (2008), *Language Awareness und bilingualer Unterricht. Eine komparative Studie*. Frankfurt/ Main: Peter Lang.
- Ganzeboom, Harry B. G.; De Graaf, Paul M.; Treimann, Donald J. & De Leeuw, Jan (1992), A standard international socio-economic index of occupational status. *Social Science Research* 25, 201-239.
- Heller, Kurt A. & Perleth, Christoph (2000), *Kognitiver Fähigkeitstest für 4. - 12. Klassen. Revision (KFT 4-12+R)*. Göttingen: Hogrefe.
- Helmke, Andreas & Schrader, Friedrich-Wilhelm (2001), Determinanten der Schulleistung. In: Rost, Detlef H. (Hrsg.) (2001), *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz PVU, 81-90.
- Hollm, Jan; Hüttermann, Armin; Keßler, Jörg-U; Schlemminger, Gérald & Ade-Thurow, Benjamin (Hrsg.) (2013), *Bilinguales Lehren und Lernen in der Sekundarstufe I. Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik.

- Kollenrott, Anne I. (2008), *Sichtweisen auf deutsch-englisch bilingualen Geschichtsunterricht. Eine empirische Studie mit Fokus auf interkulturelles Lernen*. Frankfurt/ Main: Peter Lang.
- Köller, Olaf; Knigge, Michel & Tesch, Bernd (Hrsg.) (2010), *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Köller, Olaf; Leucht, Michael & Pant, Hans A. (2012), Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I. *Unterrichtswissenschaft* 40: 4, 334-350.
- Lamsfuß-Schenk, Stefanie (2008), *Fremdverstehen im bilingualen Geschichtsunterricht. Eine Fallstudie*. Frankfurt/ Main: Peter Lang.
- Lasagabaster, David & Sierra, Juan M. (2009), Language attitudes in CLIL and traditional EFL classes. *International CLIL Research Journal* 2: 1, 4-17. [<http://www.icrj.eu/12/article1.html>].
- Madrid, Daniel (2011), Monolingual and bilingual students' competence in social sciences. In: Madrid, Daniel & Hughes, Stephen (Hrsg.) (2011), *Studies in bilingual education*. Bern: Peter Lang, 195-222.
- Marsh, Herbert W. (1990), The structure of academic self-concept: The Marsh/ Shavelson model. *Journal of Educational Psychology* 82: 4, 623-636.
- Muthén, Linda K. & Muthén, Bengt O. (1998-2012), *Mplus user's guide*. Los Angeles: Muthén & Muthén.
- Nold, Günter; Hartig, Johannes; Hinz, Silke & Rossa, Henning (2008), Klassen mit bilinguaem Sachfachunterricht: Englisch als Arbeitssprache. In: DESI-Konsortium (Hrsg.) (2008), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz, 451-457.
- Passon, Peter (2007), *Evaluation von Fachlernen und Sprachlichkeit im Kontext bilingualer Bildung*. Diplomarbeit. Universität Osnabrück. [<http://www.home.uni-osnabrueck.de/hvollmer/Diplomarbeit-Peter-Passon.pdf>].
- Pérez-Cañado, María L. (2012), CLIL research in Europe: past, present, and future. *International Journal of Bilingual Education and Bilingualism* 15: 3, 315-341.
- Raudenbush, Stephen W. & Bryk, Anthony S. (2002), *Hierarchical linear models*. Thousand Oaks: Sage.
- Rumlich, Dominik (2013), Students' general English proficiency prior to CLIL: Empirical evidence for substantial differences between prospective CLIL and non-CLIL students in Germany. In: Breidbach, Stephan & Viebrock, Britta (Hrsg.) (2013), *Content and language integrated learning (CLIL) in Europe. Research perspectives on policy and practice*. Frankfurt/ Main: Peter Lang, 181-201.
- Rumlich, Dominik (2014), Prospective CLIL and non-CLIL students' interest in English (classes): A quasi-experimental study on German sixth-graders. In: Breeze, Ruth; Pasamar, Concepción

- M.; Saíz, Carmen L. & Sala, Cristina T. (Hrsg.) (2014), *Integration of theory and practice in CLIL*. Amsterdam: Rodopi.
- UNESCO (1997), *International Standard Classification of Education I S C E D 1997*. Paris: United Nations Educational, Scientific and Cultural Organization.
- Valentine, Jeffrey C.; DuBois, David L. & Cooper, Harris (2004), The relation between self-beliefs and academic achievement. A meta-analytic review. *Educational Psychologist* 39, 111-133.
- Wigfield, Allan L. & Eccles, Jacquelynne S. (2000), Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology* 25, 68-81.
- Zydati, Wolfgang (2007), *Deutsch-Englische Zge in Berlin (DEZIBEL). Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien*. Frankfurt/ Main: Peter Lang.

Selection Effects in Content and Language
Integrated Learning Classrooms in
German Secondary Schools

Sara Dallinger

English Department, University of Education Ludwigsburg

Kathrin Jonkmann

Center for Educational Science and Psychology, University of Tübingen
Psychology Department, University of Hagen

Jan Hollm

English Department, University of Education Ludwigsburg

Abstract

Despite its high popularity throughout Europe, Content and Language Integrated Learning (CLIL) has recently received increased criticism. It has been argued that CLIL programmes are highly selective, i.e., the students possess more favourable learning prerequisites than their monolingually taught peers. The present COMBIH study contributes to this debate by investigating possible selection effects in German CLIL secondary schools. We assessed the basic cognitive skills, family background, achievement, and motivation of 1,362 CLIL and non-CLIL eighth-graders in the soon-to-be bilingually taught content subject of History. In addition, data on students' foreign language achievement and motivation were analysed with regard to whether differences in performance would diminish after students' learning prerequisites and foreign language motivation were taken into account. Results indicated statistically significant selection effects in favour of CLIL students regarding their verbal cognitive skills, parents' education, socio-economic status, cultural capital, achievement and motivation in History. Also, CLIL students showed substantial advantages in foreign language achievement and motivation. Even when selection effects and foreign language motivation of CLIL students were taken into consideration, they still outperformed their monolingually taught peers by one to almost two school years. These results can lay the foundation for an evaluation of the effectiveness of CLIL programmes.

Selection Effects in Content and Language Integrated Learning Classrooms in German Secondary Schools

Introduction

‘Content and Language Integrated Learning’ (CLIL) is a specific form of bilingual teaching. Its principles are based on teaching a content subject such as Geography or History in the students’ mother tongue and in a foreign language, thus promoting both content and language learning. This educational approach has spread rapidly in the last 10 years and has been incorporated into the educational systems of most European countries (Pérez-Cañado 2012, 319). In German political debate, CLIL is also acknowledged as an effective high-quality teaching style and has therefore been widely implemented all over the country (KMK 2013). This development has been accompanied by research that has shown that stakeholders such as students (Zydati 2007), teachers (Kollenrott 2008), the school administration (Hollm et al. 2013), and parents (Dallinger 2013) support CLIL and consider it to be highly effective. Furthermore, research focussing on students’ competences suggests that CLIL students are more competent in their foreign language skills than their monolingually taught peers (Kller, Leucht, and Pant 2012; Nold et al. 2008) and achieve similar if not better results in the bilingually taught content subject (Madrid 2011; Alonso, Grisalea, and Campo 2008). They also appear to be highly motivated (Rumlich 2014; Abendroth-Timmer 2007). In recent years, however, more critical views on CLIL have come about (Bonnet 2012; Pérez-Cañado 2012; Zydati 2012). Bruton (2011), for example, argues that previous research has neglected to consider students’ characteristics prior to the beginning of CLIL. He points out that CLIL programmes are highly selective and that, consequently, CLIL students possess more beneficial learning prerequisites than their monolingually taught peers. If this is true, then cross-sectional studies seem to have systematically overestimated the effect of CLIL (Rumlich 2014, 83). Accordingly, Bonnet argues that ‘there are strong indications from evaluation research [...] that social selection plays an important role in German CLIL. Therefore, high outcomes in CLIL groups might well be ascribed to factors other than classroom practice’ (2012, 67). The present COMBIH study (‘Competences and Motivation in Bilingual Instruction in History’) contributes further empirical evidence to aid in clearing up this debate. In a large sample of 1,362 German students from 58 classes, we investigated within-school selection effects of CLIL programmes with respect to differences between CLIL and non-CLIL students’ learning prerequisites at the onset of CLIL, such as basic verbal and non-verbal cognitive abilities, family background, and motivation and achievement in the prototypical CLIL subject of History. Moreover, we assessed differences between CLIL and non-CLIL students in English achievement and motivation over the course of the CLIL programme and sought to determine whether the observable achievement differences in English would diminish after differences in learning prerequisites, i.e., selection effects, had been taken into account.

Selection of CLIL students in German secondary schools

Potential differences in learning prerequisites between CLIL and monolingually taught students originate primarily in the selection processes of schools that offer an established CLIL programme. Such schools provide CLIL as one branch (next to other branches such as science, music, or sports) that can be chosen by students at the beginning of grade five (at age 10) after primary school. First, there are information sessions in which school representatives underline the challenges that children will have to face when attending a CLIL programme, such as higher cognitive demands, a larger homework load, more class hours, and exhibiting a pronounced willingness to learn (Rumlich 2013, 184; Bredenbröcker 2000, 50). Second, interested parents submit an application and are invited to an interview with school representatives. Third, the school administration selects appropriate students based on the interview and the applicants' primary school grades and references. In addition, there might be a lottery if too many suitable students apply for the CLIL programme (Zydati 2007, 70). This procedure is likely to lead to selection effects in favour of CLIL students with respect to their cognitive abilities, educational engagement, and family background. Since intelligence (Kuncel, Hezlett, and Ones 2004; Helmke and Schrader 2001), motivation (Valentine, DuBois, and Cooper 2004; Wigfield and Eccles 2000), and a supportive parental home (Maaz and Nagy 2010; McLoyd 1998) are important predictors of school success, they need to be taken into account and controlled for when comparing CLIL and non-CLIL students' scholastic achievements. In the following, we will summarise previous research on selection effects of bilingually taught students compared with their monolingually taught peers.

Empirical studies on the selection effects of CLIL programmes

Fehling (2008) investigated the basic verbal cognitive skills of 74 CLIL and 65 non-CLIL secondary school students at the beginning of grade seven, before the beginning of CLIL lessons. The results of this KFT subtest (*Kognitiver Fhigkeitstest*; Heller and Perleth 2000) showed that CLIL students performed significantly better than their non-CLIL peers, thus indicating more favourable learning prerequisites in favour of the bilingually taught group.

Kller, Leucht, and Pant (2012) investigated potential differences in CLIL and non-CLIL students' family backgrounds in a sample of 345 bilingually and 9,521 monolingually taught ninth-graders (students at the age of 15). They found no statistically significant differences in the International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, de Graaf, Treiman, and de Leeuw 1992) and the International Standard Classification of Education (ISCED-97; UNESCO 1997). Thus, students' socio-economic background did not differ between the two groups (cf. Zydati 2007). However, these results are based on a comparison between students from schools with and schools without CLIL programme. Comparisons *within* schools with a CLIL programme are expected to reveal differences in the CLIL and non-CLIL students' family backgrounds.

At present, there are no studies that have investigated selection effects concerning the soon-to-be bilingually taught content subject. Thus, no information is available about differences in achievement and motivation between CLIL and non-CLIL students before CLIL lessons in the particular subject begin. Yet, referring to the above results, it can be expected that the bilingually taught student group will benefit from their more favourable learning prerequisites in terms of better performances and higher motivation in the content subject.

With regard to English as a foreign language, there are several studies that have focussed on differences between CLIL and non-CLIL students' performance and motivation prior to the CLIL lessons. However, KESS 7 (Bos, Bonsen, and Gröhlich 2009) is the only known study that considered the additional English preparatory course in grades five and six. This course is obligatory for students in a CLIL programme and is bound to influence their English performance and motivation at the onset of CLIL lessons in grade seven. Accordingly, employing a listening comprehension and an overall English proficiency test (C-test) in a sample of 218 CLIL and 1,931 non-CLIL students, the authors showed that at the end of primary school (when students were 10), the CLIL students already possessed better English skills than the non-CLIL group (effect size for the comparison of means: $d = 0.50$). By the end of grade six, the CLIL students increased their English advantage even further compared with their monolingually taught peers ($d = 0.98$). Other studies whose data collection took place after the English preparatory course and before the beginning of the CLIL lessons also found this advantage in favour of CLIL students. In DENOCS, for example, Rumlich (2013) compared the overall English proficiency (C-test) of a sub-sample of 340 CLIL and 430 non-CLIL secondary school students at the end of grade six. The CLIL students performed better on the C-test than the non-CLIL group (effect size for the comparison of means: $d = 1.20$). In addition, a questionnaire was employed to measure students' language- and subject-related interest in English. The findings indicated that the bilingually taught students showed more interest in English than their monolingually taught peers with a medium effect size ($d = 0.50$; Rumlich 2014). Fehling (2008) replicated Rumlich's results. Besides the basic cognitive abilities (see above), she also investigated the five dimensions of language awareness (cf. James and Garrett 1992) and showed that, at the beginning of grade seven, CLIL students were already outperforming the non-CLIL group in three of four language-awareness dimensions (affective, cognitive, political).

In order to account for these substantial differences in English achievement in favour of the bilingually taught students, a few studies controlled for the CLIL and non-CLIL groups' learning prerequisites. In the study reported above by Köller, Leucht, and Pant (2012), the following covariates were considered in predicting the students' reading and listening comprehension skills: gender, mother tongue, competences in the mother tongue, immigration background, ISEI, and ISCED. They found that the CLIL students' advantage on the English tests barely diminished (only by 0.02 standard deviations), thus indicating that there were no selection effects. Thus, the CLIL students' outstanding performance appeared to be independent of learning prerequisites, and the

authors concluded that CLIL enhances the students' foreign language competences. Nold et al. (2008) obtained similar findings. An examination of the English skills of a DESI sub-sample of almost 2,000 secondary school students (grade nine, age 15) indicated that the CLIL students achieved a higher competence level than the comparison group with regard to their English listening, reading, grammar, and text reconstruction skills even though the CLIL and non-CLIL students were previously matched by gender, mother tongue, competences in the mother tongue, basic cognitive skills, and parents' profession and education. As yet, no studies have controlled for motivational aspects, such as interest in or importance of the English language.

To summarise, the few existing findings have demonstrated selection effects in favour of the CLIL students regarding their basic verbal cognitive abilities. Concerning socio-economic background, Köller, Leucht, and Pant (2012) did not find any differences between the student groups. Their analysis, however, did not explicitly compare students within schools that offered a CLIL programme. Furthermore, there is no information available about the students' prior knowledge and motivation in a soon-to-be bilingually taught content subject. Regarding English achievement and motivation, results indicated substantial advantages in favour of CLIL students. Only a few studies, however, have controlled for learning prerequisites when estimating these advantages. Köller, Leucht, and Pant (2012) found no noteworthy selection effect. Their study and Nold et al.'s (2008) showed that CLIL students' English performance was still statistically significantly better than the performance of the non-CLIL group even though covariates were considered. Additional aspects such as motivation and content-subject-related matters have been neglected so far.

Research questions

Given the empirical evidence presented above, there is a need for more solid research to comprehensively examine the within-school selection effects of CLIL programmes. This is true not only for students' basic cognitive skills and family background, but also for other indicators of learning prerequisites such as motivation and achievement in a soon-to-be bilingually taught content subject. Furthermore, this information should be used to control for differences between CLIL and monolingually taught students over the course of CLIL programmes to gain tentative insights into the programmes' effectiveness. This is one research focus of COMBIH. The present study addressed three research questions. Research question 1 asked whether selection effects favouring CLIL students could be observed with regard to their general cognitive abilities (verbal and figural cognitive skills), family background (immigration background, parents' education, socio-economic status, cultural capital), and motivation (interest, importance) and performance (standardised achievement test) in the content subject History. We expected that CLIL students would show better basic cognitive skills than their monolingually taught peers and, moreover, a more favourable family background since all group comparisons involved schools with a CLIL programme. As a consequence of these expected differences in learning prerequisites, we expected that CLIL students would also show better performances and higher motivation in History. Research question 2 asked

about the extent to which CLIL and non-CLIL students would differ in their English motivation and performance over the course of a CLIL programme. We expected that the CLIL students would exhibit substantial advantages. The third research question asked whether the English achievement differences would diminish after controlling for the selection effects. In additional analyses, we tested whether these differences would decline further after motivational differences in English were taken into account. We expected that the CLIL students' advantage in English achievement would diminish to some extent but would still be statistically significant compared with their monolingually taught peers.

Method

Sample

The sample was comprised of 1,362 eighth graders from 28 academic-track secondary schools ('*Gymnasien*') in the federal state of Baden-Wuerttemberg (49.3% girls; $M_{age} = 13.5$, $SD_{age} = 0.51$). 69.3% of the students reported that at least one parent had qualified for college education (i.e., obtained the *Abitur* certificate). 28% of the students reported having an immigration background (i.e., at least one parent was not born in Germany). The sample consisted of two groups. The first group was comprised of 703 students from 30 classes who took part in a CLIL programme. In grades five and six, they attended two additional English lessons per week. From grade seven on, at least one subject was taught bilingually throughout the school year. In grade seven, this subject was Geography; in grade eight, students took part in bilingual History lessons for the first time. They had already attended regular History lessons taught in their mother tongue in grades six and seven. The second group was comprised of 659 students from 28 classes. The students were from the same schools as the CLIL students but participated in a different branch. They participated in regular English lessons beginning in grade five and in regular History lessons taught in their mother tongue beginning in grade six.

Procedure

The data were collected in autumn 2012 at the beginning of grade eight. Beforehand, permission was obtained from the Ministry of Education and Cultural Affairs of Baden-Wuerttemberg to conduct the study. Student participation was voluntary, and written consent was obtained from their parents. Excluding breaks, the data collection took 135 minutes and was conducted during regular class time by research assistants with the help of a standardised manual. Students took several performance tests and then filled out a questionnaire. They worked on the material by themselves in a paper-and-pencil format.

Instruments

Cognitive abilities. Students completed the figural and verbal analogies subtests of the KFT (*Kognitiver Fähigkeitstest*; Heller and Perleth 2000). The tests asked the students to detect two matching figures or a matching word pair on the basis of the selection of several distracting figures or words. Parallel test versions were employed to prevent students from copying their neighbours' results. The figural test was comprised of 25 items with a reliability of $\alpha = .91$ for test version A and $\alpha = .90$ for test version B. The verbal test was comprised of 20 items with $\alpha = .69$ for version A and $\alpha = .72$ for version B.

Family background. The items measuring the students' family background were adapted from large-scale studies such as PISA (Klieme et al. 2010). On the basis of students' information about their parents' occupation and education, the International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom et al. 1992; HISEI for the higher ISEI of the two parents) was determined for each student, ranging between 11 (very low socio-economic status) and 88 (very high socio-economic status). To measure cultural capital in the family, we asked for the number of books per household. Responses were given on a six-point scale ranging from '0 to 10 books' to 'more than 500 books'. Finally, we collected information about students' gender, immigration background (at least one parent was not born in Germany), and parents' educational degree (at least one parent had qualified for college education).

Motivation. Motivation in History and motivation in English were each assessed with two indicators that were based on the value component of the expectancy-value model of achievement motivation by Eccles (e.g., Wigfield and Eccles 2000). We assessed subject-specific intrinsic value (five items for each subject; History: $\alpha = .82$; English: $\alpha = .72$; sample item: 'History/English is one of my favourite subjects') and attainment value (three items for each subject; History: $\alpha = .80$; English: $\alpha = .83$; sample item: 'I think History/English is a very important school subject'). Responses were given on a four-point Likert-type scale ranging from one ('completely disagree') to four ('completely agree').

History achievement. Students' knowledge of history was tested with a cloze test and multiple-choice questions totalling 25 items ($\alpha = .72$). The test measured the students' content knowledge from the previous grade (grade seven) and was self-developed on the basis of the state's History curriculum. The test was administered in German. The measurement properties and validity of the test were established in a pilot study with a sample of 222 eighth graders.

English achievement. Students' global English performance was assessed with a self-developed overall proficiency test (C-test). Employing C-tests to measure overall English performance is a standard procedure that has been applied in many studies, such as DENOCS (Rumlich 2013), DESI (DESI-Konsortium 2008), or DEZIBEL (Zydati 2007). A C-test consists of several separate short texts that cover different topics. In these passages, the second half of every

other word is missing (Grotjahn 2011). For the present analyses, the C-test was comprised of three subtexts with a total of 70 items and a reliability of $\alpha = .89$.

Finally, English listening comprehension was assessed with items from the IQB (*Institut für Qualitätsentwicklung im Bildungswesen*, Berlin/ Germany; cf. Köller, Knigge, and Tesch 2010). It consisted of six tasks with a total of 27 items ($\alpha = .73$). The test contained open and closed questions. Students were asked to listen to the texts and to answer the questions within a time span that ranged from five to 30 seconds.

Analyses

As COMBIH was conducted in a classroom setting, the data comprised a hierarchical structure in which individuals were nested within classes. Since two students from the same class are more similar to one another, on average, than two randomly selected students from the population, the assumption of total independence of observations was violated. To take this into account, all analyses were computed in Mplus 7.1 (Muthén and Muthén 1998–2012) using the command ‘type = complex’. Thus, the standard errors of the regression coefficients were adjusted to prevent their underestimation as well as to prevent the significance testing from being too liberal (Raudenbush and Bryk 2002). The percentage of missing values ranged from 0% (gender) to 17% (HISEI). To deal with missing data, the missing value estimator in Mplus 7.1 was used. Mplus applies a model-based approach, which builds on full information maximum likelihood (FIML) estimation.

Results

Research question 1: Differences between CLIL and non-CLIL students in basic cognitive abilities, family background, and motivation and competences in History

Table 1 displays group means and standard deviations for gender, family background, general cognitive abilities, and achievement and motivation in History for CLIL and non-CLIL students. The results of the significance tests for the group comparisons are also displayed.

Although there were slightly more girls in the CLIL group, the two groups did not exhibit statistically significant differences with respect to the distribution of gender (Table 1). The same was true for the distribution of students with immigration backgrounds. Their percentage was higher in the CLIL group than in the non-CLIL group, but the difference was not statistically significant. However, regarding their parents’ education, socio-economic status, and the number of books per household, CLIL students showed significantly higher values than their monolingually taught peers. These differences amounted to 0.20 and 0.17 standard deviations. Regarding the figural cognitive skills test, CLIL students achieved slightly better results than their non-CLIL peers; however, this advantage was not statistically significant. By contrast, CLIL students achieved markedly better results on the verbal cognitive skills test with an advantage of over one third of a standard deviation compared with the non-CLIL group. On the history test, the CLIL group outperformed their peers by

0.29 standard deviations, thus indicating a performance advantage of approximately three quarters of a school year in favour of the CLIL students (cf. Köller, Knigge, and Tesch 2010). In addition, for the ‘attainment value’ variable, the CLIL students were also more motivated in History than their monolingually taught peers. Thus, the present results supported our expectations with regard to which differences in basic verbal cognitive skills, family background (parents’ education, socio-economic status, number of books per household), and motivation and competences in History would favour the CLIL students.

Tab. 1: Differences in gender, family background, general cognitive abilities, and achievement and motivation in History between CLIL and non-CLIL students

| | CLIL | | non-CLIL | | CLIL vs. non-CLIL | |
|-----------------------------|----------|-----------|----------|-----------|-------------------|----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>d</i> | <i>p</i> |
| Girls | 50.1% | - | 48.6% | - | - | ns |
| With immigration background | 30.5% | - | 25.5% | - | - | ns |
| Parents’ education high | 73.9% | - | 64.2% | - | - | * |
| HISEI | 65.42 | 17.72 | 61.87 | 18.70 | .20 | ** |
| Cultural capital (books) | 4.84 | 1.17 | 4.63 | 1.23 | .17 | * |
| KFT figural | 17.51 | 6.25 | 16.55 | 6.43 | .15 | ns |
| KFT verbal | 13.88 | 3.39 | 12.62 | 3.46 | .37 | *** |
| History test | 11.40 | 3.94 | 10.20 | 4.33 | .29 | * |
| Intrinsic value History | 2.26 | 0.69 | 2.16 | 0.76 | .13 | ns |
| Attainment value History | 2.77 | 0.66 | 2.59 | 0.69 | .25 | *** |

Note: CLIL = CLIL student group, non-CLIL = non-CLIL student group; *M* = mean, *SD* = standard deviation; HISEI = International Socio-Economic Index of Occupational Status; KFT = basic cognitive skills test.

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

Research question 2: Differences between CLIL and non-CLIL students in English motivation and competences

Table 2 displays group means and standard deviations for English achievement and motivation for CLIL and non-CLIL students. The significance test results for the group comparisons are also displayed.

The findings indicated that the CLIL students performed significantly better in listening comprehension and on the C-test than the non-CLIL group. The differences reflected large effects totalling 0.71 and 0.93 standard deviations. Thus, the CLIL students outperformed their monolingually taught peers by a range of almost three quarters to one standard deviation, which equals a performance advantage of up to two school years. Regarding the two motivational variables, the CLIL students showed higher values than the non-CLIL group with group mean differences ranging between one third and one half of a standard deviation. In conclusion, the CLIL students showed substantial advantages on the C-test, listening comprehension, and motivation in English, as expected from research question 2.

Tab. 2: Differences in English achievement and motivation between CLIL and non-CLIL students

| | CLIL | | non-CLIL | | CLIL vs. non-CLIL | |
|--------------------------|----------|-----------|----------|-----------|-------------------|----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>d</i> | <i>p</i> |
| C-test | 44.36 | 8.55 | 35.60 | 8.52 | .93 | *** |
| Listening comprehension | 22.53 | 2.53 | 20.31 | 3.39 | .71 | *** |
| Intrinsic value English | 2.98 | 0.62 | 2.72 | 0.67 | .40 | *** |
| Attainment value English | 3.53 | 0.54 | 3.30 | 0.63 | .39 | *** |

Note: CLIL = CLIL student group, non-CLIL = non-CLIL student group; *M* = mean, *SD* = standard deviation. * $p < .05$, ** $p < .01$, *** $p < .001$.

Research question 3: English achievement differences between CLIL and non-CLIL students when selection effects and English motivation were taken into consideration

Two multiple regression models were estimated to take CLIL selection effects and English motivation into account when examining English achievement differences between CLIL and non-CLIL students on the listening comprehension test and the C-test. Model 1 considered ten covariates: students' gender, immigration background, parents' education, socio-economic status, number of books per household, figural and verbal cognitive skills tests, History achievement, and intrinsic value and attainment value attributed to History. Model 2 was expanded by adding the two English motivational variables. In both models, students' affiliation with the CLIL or non-CLIL group was included as another independent variable. Table 3 displays the standardised regression coefficients, standard errors, and the analyses that were applied to test for statistical significance.

The table shows that the statistically significant effects of the covariates on English achievement were rather small, ranging between 0.06 and 0.24 standard deviations. All other variables being equal, girls exhibited higher C-test performances than boys. Similarly, students with immigration backgrounds achieved better test results than students whose parents were born in Germany. Furthermore, higher socio-economic status and higher figural and verbal cognitive abilities were also significant predictors of C-test performance. Regarding listening comprehension, higher figural and verbal cognitive abilities, better History achievement, and less intrinsic value attributed to History significantly predicted test performance when all other variables were held equal.

Model 1 indicated that controlling for potential group differences between monolingually and bilingually taught students in gender, learning prerequisites, and motivation and achievement in History resulted in a substantial reduction in the CLIL students' advantages in English performance. On the C-test, it diminished from 0.93 to 0.74, and for listening comprehension, from 0.71 to 0.54 standard deviations. Thus, controlling for selection effects substantially reduced the CLIL students' advantage by 0.19 and 0.17 standard deviations. Additionally considering the English motivational variables in Model 2 resulted in another reduction in the CLIL students' advantage in English performance of 0.74 to 0.68 standard deviations on the C-test and 0.54 to 0.51 standard deviations in listening comprehension. Thus, the non-CLIL students' performance disadvantage was slightly

reduced but was still substantial and statistically significant. In conclusion, even when taking selection effects and English motivation into consideration, the CLIL group still outperformed their monolingually taught peers on the English tests by one and a quarter to one and three quarters school years.

Tab. 3: English achievement differences between CLIL and non-CLIL students when selection effects were taken into consideration

| | C-test | | | | Listening comprehension test | | | |
|----------------------------------|---------|------------|---------|------------|------------------------------|------------|---------|------------|
| | Model 1 | | Model 2 | | Model 1 | | Model 2 | |
| | β | (SE) | β | (SE) | β | (SE) | β | (SE) |
| Reference group | 0.26 | ***(0.07) | 0.28 | ***(0.07) | 0.20 | ** (0.07) | 0.21 | ** (0.07) |
| Non-CLIL (1 = non-CLIL) | -0.74 | ***(0.07) | -0.68 | ***(0.06) | -0.54 | *** (0.07) | -0.51 | *** (0.07) |
| Gender (1 = f) | 0.28 | *** (0.05) | 0.24 | *** (0.05) | 0.10 | (0.07) | 0.07 | (0.07) |
| Immigration background (1 = yes) | 0.15 | ** (0.05) | 0.12 | * (0.05) | 0.10 | * (0.05) | 0.09 | (0.05) |
| Parents' education (1= high) | -0.13 | * (0.06) | -0.16 | ** (0.06) | -0.03 | (0.07) | -0.04 | (0.07) |
| HISEI | 0.05 | (0.03) | 0.06 | * (0.03) | -0.01 | (0.03) | 0.00 | (0.03) |
| Cultural capital (books) | 0.04 | (0.03) | 0.03 | (0.03) | 0.03 | (0.03) | 0.03 | (0.03) |
| KFT figural | 0.09 | *** (0.03) | 0.09 | *** (0.03) | 0.09 | *** (0.03) | 0.10 | *** (0.03) |
| KFT verbal | 0.20 | *** (0.03) | 0.19 | *** (0.03) | 0.19 | *** (0.03) | 0.18 | *** (0.03) |
| History test | 0.19 | *** (0.03) | 0.19 | *** (0.03) | 0.20 | *** (0.04) | 0.20 | *** (0.04) |
| Intrinsic value History | -0.08 | * (0.03) | -0.07 | * (0.03) | -0.14 | *** (0.04) | -0.13 | *** (0.04) |
| Attainment value History | 0.04 | (0.03) | -0.01 | (0.03) | 0.07 | * (0.04) | 0.05 | (0.04) |
| Intrinsic value English | | | 0.14 | *** (0.03) | | | 0.07 | (0.04) |
| Attainment value English | | | 0.06 | (0.03) | | | 0.04 | (0.03) |
| R ² | 0.34 | | 0.37 | | 0.24 | | 0.25 | |

Note: β = standardised regression coefficient, SE = standard error; CLIL = CLIL student group, non-CLIL = non-CLIL student group; HISEI = International Socio-Economic Index of Occupational Status, KFT = basic cognitive skills test.

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

Summary and discussion

The aim of this article was to investigate the within-school selection effects of CLIL students compared with their monolingually taught peers and to control for these selection effects when examining differences in English achievement between CLIL and non-CLIL students. For research question 1, the results indicated more favourable learning prerequisites for CLIL students with regard to their verbal cognitive skills, parents' education, socio-economic status, and cultural capital. Small group differences on the figural cognitive skills test and the distributions of gender and immigration background were not statistically significant. In addition, the selection effect was demonstrated in the prototypical CLIL subject of History. At the beginning of bilingually taught History lessons, the CLIL students already showed higher motivation and better test performances in this subject than the non-CLIL group. For research question 2, the findings again indicated substantial differences in English motivation and achievement in favour of the CLIL students. Furthermore, the results for

research question 3 showed that the CLIL group's advantage in English performance was substantially diminished when the selection and English motivation effects that we reported above were controlled for. Yet, the CLIL group's advantage was still statistically significant and amounted to an advantage of one and a quarter to one and three quarters school years.

The present COMBIH study replicated Fehling's (2008) results in which CLIL students showed better verbal cognitive skills than their monolingually taught peers. In addition, the present study found no statistically significant differences in figural cognitive abilities. Furthermore, the present results indicated that the CLIL students possessed a higher socio-economic status than their monolingually taught peers from the same schools. These findings contradict Köller, Leucht, and Pant's (2012) previous research, which, however, was based on comparisons between students from schools with and without a CLIL programme. Nevertheless, the present COMBIH results confirm the previously suggested within-school selection effect of family background since there was more cultural capital in CLIL students' households, and their parents' educational degree was higher than that of the non-CLIL group. Moreover, the present regression analyses showed that the family background variables were in fact related to students' performance in English.

Concerning the bilingually taught content subject, previous studies have shown that CLIL students achieve similar or slightly better results than their monolingually taught peers (Madrid 2011; Alonso, Grisaleña, and Campo 2008). The present findings, however, took into account student achievement and motivation in the subject of History before it was taught bilingually in order to account for potential selection effects. Our results indicated that CLIL students outperformed their peers by approximately three quarters of a school year, i.e., they possessed more prior knowledge of History than the comparison group. Also, they were more motivated in History than the non-CLIL students at least with respect to the importance of the subject. These findings supported our expectations and could most likely be ascribed to the CLIL students' more favourable learning prerequisites given that cognitive abilities (Kuncel, Hezlett, and Ones 2004; Helmke and Schrader 2001) and a supportive parental home (Maaz and Nagy 2010; McLoyd 1998), for example, are closely linked to school success. The results of the regression analyses underlined this prediction. Another reason for the CLIL students' advantages could be that History is a language-based subject and is therefore in line with their interests and skills in (foreign) languages. The fact that the CLIL students did not attribute a higher intrinsic value to History could be explained by the contents of the subject because participation in a CLIL programme does not automatically imply an increased interest in historical events. Yet, altogether, we can conclude that CLIL students are endowed with more favourable learning prerequisites than their monolingually taught peers from the outset, a finding that appears to be a consequence of the selectivity of CLIL programmes.

The outstanding performance of CLIL students on English tests compared with their peers has been a hitherto frequently replicated result (Rumlich 2013; Köller, Leucht, and Pant 2012; Nold et al. 2008) and is not surprising in general considering their enhanced exposure to this foreign

language. In the present study, an attempt was made to explain these differences between CLIL and non-CLIL students by controlling for the previously found selection effects in their general cognitive abilities, family background, and motivation and competences in History. A multiple regression model showed that the CLIL students' advantages were considerably smaller after such selection effects were controlled for. Thus, their more favourable learning prerequisites play an important role in explaining the bilingually taught students' outstanding English performance and should therefore be considered in all future research investigating differences in CLIL and non-CLIL students' foreign language achievement. Nevertheless, there are still substantial differences in performance left to be accounted for. It is likely that the CLIL students' high English motivation as already reported by Rumlich (2014) and replicated in COMBIH can account for this to some degree. One must assume, however, that the CLIL students' English motivation at the beginning of grade eight is already strongly influenced by their CLIL programme experiences in grades five to seven. Therefore, we subsequently estimated another regression model that included two English motivational variables. Our results showed that the CLIL students' advantage diminished again, yet there was still a substantial performance advantage that amounted to one and a quarter to one and three quarters school years. This unaccounted-for variance might provide information about the effectiveness of a CLIL programme, i.e., about the contribution of the CLIL programme to CLIL students' English competences. Yet, it would require longitudinal studies to control for the presently reported selection effects and advantages in English motivation and performance when reporting the development of competences to optimally isolate the effect of CLIL. Moreover, these kind of studies would need to consider additional aspects that might further diminish differences in CLIL and non-CLIL students' achievement. Such aspects may consist of the additional preparatory English lessons the CLIL students receive in grades five and six, as suggested by Bos, Bensen, and Gröhlich (2009). Other aspects may consist of students' extra-curricular contact with English as a foreign language and potential selection effects by CLIL teachers in comparison with non-CLIL teachers.

In summary, the present results indicate the presence of strong selection effects in CLIL programmes. CLIL students possess more favourable learning prerequisites compared with their monolingually taught peers with respect to their verbal cognitive abilities, family background, and motivation and achievement in the prototypical content subject of History. Also, CLIL students show substantially higher motivation and better test results in English over the course of the CLIL programme than the non-CLIL group. This advantage in English performance is clearly diminished but not completely accounted for after selection effects and English motivation are controlled for. Ideally, future research will take more aspects into account, such as potential selection effects from the teaching staff or students' extra-curricular contact with English as a foreign language. Future studies should also consider these selection effects and integrate them into a longitudinal study, which would then lay the foundation for an evaluation of the effectiveness of CLIL programmes.

References

- Abendroth-Timmer, Dagmar. 2007. *Akzeptanz und Motivation: Empirische Ansätze zur Erforschung des unterrichtlichen Einsatzes von bilingualen und mehrsprachigen Modulen [Acceptance and Motivation: Empirical Approaches to Investigating the Usage of Bilingual and Multilingual Modules in the Classroom]*. Frankfurt am Main: Lang.
- Alonso, Esmeralda, Jesús Grisaleña, and Alejandro Campo. 2008. "Plurilingual Education in Secondary Schools: Analysis of Results." *International CLIL Research Journal* 1 (1): 36–49. <http://www.icrj.eu/11/article3.html/>.
- Bonnet, Andreas. 2012. "Towards an Evidence Base for CLIL: How to Integrate Qualitative and Quantitative as Well as Process, Product and Participant Perspectives in CLIL Research." *International CLIL Research Journal* 4 (1): 66–78. <http://www.icrj.eu/14/article7.html/>.
- Bos, Wilfried, Martin Bosen, and Carola Gröhlich. 2009. *KESS 7: Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7 [KESS 7: Competences and Attitudes of Students from Hamburg Schools at the Beginning of Form 7]*. Münster: Waxmann.
- Bredenbröcker, Winfried. 2000. *Förderung der fremdsprachlichen Kompetenz durch bilingualen Unterricht: Empirische Untersuchungen [Promoting Foreign Language Competences with the Help of Content and Language Integrated Learning: Empirical Investigations]*. Frankfurt am Main: Lang.
- Bruton, Anthony. 2011. "Is CLIL so Beneficial, or Just Selective? Re-evaluating Some of the Research." *System* 39 (4): 523–532.
- Dallinger, Sara. 2013. "Die Einstellung der Eltern zum bilingualen Sachfachunterricht: Eine quantitative Studie an Realschulen in Baden-Württemberg [Attitude of Parents Towards Content and Language Integrated Learning: A Quantitative Study at Realschulen in Baden-Wuerttemberg]." In *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation [Bilingual Teaching and Learning in Secondary Education: Language, Content Subject and School Organization]*, edited by Jan Hollm, Armin Hüttermann, Jörg-U. Keßler, Gérald Schlemminger, and Benjamin Ade-Thurow, 268–286. Landau: Empirische Pädagogik.
- DESI-Konsortium, eds. 2008. *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie [Lessons and Acquisition of Competences in German and English: Results from the DESI Study]*. Weinheim: Beltz.
- Fehling, Sylvia. 2008. *Language Awareness und bilingualer Unterricht: Eine komparative Studie [Language Awareness and Content and Language Integrated Learning: A Comparative Study]*. Frankfurt am Main: Lang.

- Ganzeboom, Harry, Paul de Graaf, Donald Treimann, and Jan de Leeuw. 1992. "A Standard International Socio-Economic Index of Occupational Status." *Social Science Research* 21 (1): 1–56.
- Grotjahn, Rüdiger. 2011. "C-Tests: Aspekte der Validität [C-tests: Aspects of Validity]." *Deutsch als Fremdsprache* 48 (3): 131–137.
- Heller, Kurt, and Christoph Perleth. 2000. *Kognitiver Fähigkeitstest für 4. - 12. Klassen: Revision (KFT 4-12+R) [Cognitive Abilities Test for 4th - 12th Graders: Revision (KFT 4-12+R)]*. Göttingen: Hogrefe.
- Helmke, Andreas, and Friedrich-Wilhelm Schrader. 2001. "Determinanten der Schulleistung [Determinants of School Achievement]." In *Handwörterbuch Pädagogische Psychologie [Concise Dictionary of Pedagogical Psychology]*, edited by Detlef Rost, 81–90. Weinheim: Beltz PVU.
- Hollm, Jan, Armin Hüttermann, Jörg-U. Keßler, Gérald Schlemminger, and Benjamin Ade-Thurow, eds. 2013. *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation [Bilingual Teaching and Learning in Secondary Education: Language, Content Subject and School Organization]*. Landau: Empirische Pädagogik.
- James, Carl, and Peter Garrett. 1992. *Language Awareness in the Classroom*. London: Longman.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) 2013. "Bericht 'Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung': Beschluss der Kultusministerkonferenz vom 17.10.2013 [Report 'Concepts for Content and Language Integrated Learning – Report Based on Experiences and Suggestions for Further Developments': Resolution of the The Standing Conference of the Ministers of Education and Cultural Affairs from 17.10.2013]." http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/201_10_17-Konzepte-_bilingualer-_Unterricht.pdf.
- Klieme, Eckhard, Cordula Artelt, Johannes Hartig, Nina Jude, Olaf Köller, and Manfred Prenzel, eds. 2010. *PISA 2009*. Münster: Waxmann.
- Kollenrott, Anne. 2008. *Sichtweisen auf deutsch-englisch bilingualen Geschichtsunterricht: Eine empirische Studie mit Fokus auf interkulturelles Lernen [Perspectives on German-English Content and Language Integrated Learning in History: An Empirical Study Focussing on Intercultural Learning]*. Frankfurt am Main: Lang.
- Köller, Olaf, Michel Knigge, and Bernd Tesch, eds. 2010. *Sprachliche Kompetenzen im Ländervergleich [Comparison of Linguistic Competences in German States]*. Münster: Waxmann.
- Köller, Olaf, Michael Leucht, and Hans Pant. 2012. "Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I [Effects of Bilingual Instruction on English Achievement in Lower Secondary Schools]." *Unterrichtswissenschaft* 4 (4): 334–350.

- Kuncel, Nathan; Sarah Hezlett, and Deniz Ones. 2004. "Academic Performance, Careerpotential, Creativity, and Job Performance: Can One Construct Predict them all?" *Journal of Personality and Social Psychology* 86: 148–161.
- Maaz, Kai, and Gabriel Nagy. 2010. "Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte [Transition from Primary School to Consecutive Secondary Schools: Definition, Specification and Quantification of Primary and Secondary Effects of Origin]." *Zeitschrift für Erziehungswissenschaft Sonderheft* 12: 153–182.
- Madrid, Daniel. 2011. "Monolingual and Bilingual Students' Competence in Social Sciences." In *Studies in Bilingual Education*, edited by Daniel Madrid, and Stephen Hughes, 195–222. Bern: Lang.
- McLoyd, Vonnie C. 1998. "Socioeconomic Disadvantage and Child Development." *American Psychologist* 53 (2): 185–204.
- Muthén, Linda, and Bengt Muthén. 1998–2012. *Mplus User's Guide*. Los Angeles: Muthén & Muthén.
- Nold, Günter, Johannes Hartig, Silke Hinz, and Henning Rossa. 2008. "Klassen mit bilinguaem Sachfachunterricht: Englisch als Arbeitssprache [Classes with Content and Language Integrated Learning: English as Language of Communication]." In *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie [Lessons and Acquisition of Competences in German and English: Results from the DESI Study]*, edited by DESI-Konsortium, 451–457. Weinheim: Beltz.
- Pérez-Cañado, María. 2012. "CLIL Research in Europe: Past, Present, and Future." *International Journal of Bilingual Education and Bilingualism* 15 (3): 315–341.
- Raudenbush, Stephen, and Anthony Bryk. 2002. *Hierarchical Linear Models*. Thousand Oaks: Sage.
- Rumlich, Dominik. 2013. "Students' General English Proficiency Prior to CLIL: Empirical Evidence for Substantial Differences between Prospective CLIL and non-CLIL Students in Germany." In *Content and Language Integrated Learning (CLIL) in Europe: Research Perspectives on Policy and Practice*, edited by Stephan Breidbach, and Britta Viebrock, 181–201. Frankfurt am Main: Lang.
- Rumlich, Dominik. 2014. "Prospective CLIL and non-CLIL Students' Interest in English (Classes): A Quasi-Experimental Study on German Sixth-Graders." In *Integration of Theory and Practice in CLIL*, edited by Ruth Breeze, Concepción Pasamar, Carmen Saíz, and Cristina Sala, 75–95. Amsterdam: Rodopi.
- UNESCO. 1997. *International Standard Classification of Education ISCED 1997*. Paris: United Nations Educational, Scientific and Cultural Organization.

- Valentine, Jeffrey, David DuBois, and Harris Cooper. 2004. "The Relation between Self-Beliefs and Academic Achievement: A Meta-Analytic Review." *Educational Psychologist* 39: 111–133.
- Wigfield, Allan, and Jacquelynne Eccles. 2000. "Expectancy-Value Theory of Achievement Motivation." *Contemporary Educational Psychology* 25: 68–81.
- Zydati, Wolfgang. 2007. *Deutsch-Englische Zge in Berlin (DEZIBEL): Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien [German-English Profiles in Berlin (DEZIBEL): Evaluation of Content and Language Integrated Learning at Gymnasien]*. Frankfurt am Main: Lang.
- Zydati, Wolfgang. 2012. "Linguistic Thresholds in the CLIL Classroom? The Threshold Hypothesis Revisited." *International CLIL Research Journal* 4 (1): 17–28. <http://www.icrj.eu/14/article2.html/>.

The Effect of Content and Language Integrated
Learning on Students' English and
History Competences

– Killing Two Birds With One Stone?

Sara Dallinger

English Department, University of Education Ludwigsburg

Kathrin Jonkmann

Center for Educational Science and Psychology, University of Tübingen
Psychology Department, University of Hagen

Jan Hollm

English Department, University of Education Ludwigsburg

Christiane Fiege

Institute of Psychology, University of Tübingen

Dallinger, Sara; Jonkmann, Kathrin; Hollm, Jan & Fiege, Christiane (2014). *The Effect of Content and Language Integrated Learning on Students' English and History Competences – Killing Two Birds With One Stone?* Manuscript submitted for publication.
(The citation conforms to the journal's style guidelines.)

Abstract

By using cross-sectional designs and failing to consider confounding variables, previous research has overestimated the effects of Content and Language Integrated Learning (CLIL) on the development of students' foreign language competences. Furthermore, not many longitudinal studies have addressed the effects of CLIL on the content subject that is taught bilingually. Therefore, the present COMBIH study investigated the competence development of 1,806 German CLIL and non-CLIL eighth-graders in English and History while considering students' prior achievement, gender, immigration background, basic cognitive skills, socioeconomic status, motivation, instructional quality, and teacher characteristics. Results showed greater increases in English listening comprehension in CLIL than non-CLIL classrooms but not in general English skills. In History, the increases in subject knowledge were comparable. The findings confirm previous, differential findings for English and show that CLIL is not detrimental to content-subject achievement. These results might provide a foundation for future decisions about the expansion of CLIL programmes.

The Effect of Content and Language Integrated Learning on Students' English and History Competences - Killing Two Birds With One Stone?

Introduction

Content and Language Integrated Learning (CLIL) is a specific approach to bilingual teaching and learning that has spread across Europe in the last two decades and is steadily gaining popularity (Pérez-Cañado, 2012). It refers to an educational setting in which both the students' mother tongue and a foreign language are used when teaching a content subject to consequently foster both content and language learning (Eurydice, 2006). Previous studies have predominantly focused on its effects on the foreign language and have shown that students' foreign language skills benefitted from CLIL (Köller, Leucht, & Pant, 2012; Nold, Hartig, Hinz, & Rossa, 2008). However, only a little research has been conducted on the effects of CLIL on student competences in the bilingually taught content subject. Does content learning benefit or suffer from CLIL, or does it not make a difference? The few findings have suggested that the latter might be the case (Badertscher & Bieri, 2009; Madrid, 2011), but the empirical database is still weak. Furthermore, criticisms of CLIL research have been intensifying in recent years (Bonnet, 2012; Bruton, 2011; Pérez-Cañado, 2012; Rumlich, 2014a). Rumlich (2014a), for example, argued that most cross-sectional studies seem to have overestimated the effects of CLIL on the development of student competences. Similarly, Bruton (2011) pointed out that analyses of students' learning prerequisites are frequently missing and argued that CLIL programmes cause selection effects that favour CLIL students. Accordingly, Dallinger and Jonkmann (2014) showed that CLIL students have better verbal cognitive abilities, more favourable family backgrounds, better performances, and higher motivation in a soon-to-be bilingually taught content subject from the outset. Thus, when estimating the effects of CLIL, these confounding variables need to be taken into account. Pérez-Cañado (2012) concluded that there is a "need of solid empirical research which builds in rigorous assessment of the variables under scrutiny [...] to determine whether the gains observed are truly ascribable to CLIL practice" (p. 330). There is thus a need for longitudinal studies to consider possible selection effects of CLIL programmes and other confounding factors when identifying the effects of CLIL (Costa & D'Angelo, 2011). The present "Competences and Motivation in Bilingual Instruction in History" (COMBIH) study was designed to fill this research gap. In a large sample of 1,806 German secondary school students, we investigated the development of competences in English as a foreign language and in the prototypical CLIL subject of History over the course of Grade 8. To isolate the CLIL effect as much as possible, we considered prior achievement, learning prerequisites, quality of instruction, and teacher characteristics in multilevel regression analyses.

Theoretical and Empirical Background

Theories about Effects of CLIL on Language and Content Learning

The value of CLIL programmes has been debated for years. Especially at the beginning of enhanced CLIL research, studies aimed to demonstrate CLIL's positive contribution to students' foreign language learning (Breidbach & Viebrock, 2012). This has its roots in several theoretical approaches of second language acquisition. The input hypothesis by Krashen (1985; part of the monitor model) supports the assumption that students benefit from CLIL in the development of their foreign language competences. It suggests that language is acquired with the help of a sufficient number of understandable messages (i.e., comprehensible input). The acquisition process is accelerated if the input is slightly beyond the learner's current level of competence. These conditions are found in CLIL classrooms because the students are usually exposed to a larger amount of (oral and written) foreign language input than non-CLIL students, which is typically located on a competence level that is higher than that of the students. Because CLIL teachers are usually aware of the linguistic challenges that their students meet when learning a content subject in a foreign language, the language input is frequently made more comprehensible through "foreigner talk" (i.e., simplified speech; Krashen, 1985, p. 8) and additional linguistic and nonlinguistic material (e.g., vocabulary lists, pictures).

Foreign language input plays an essential role in the interaction approach, too (Gass & Mackey, 2007). It also supports the assumption that a CLIL setting is beneficial for students' foreign-language competences. The approach suggests that apart from an appropriate amount of language input, students also need to be provided with sufficient opportunities to interact with the language and to test and adjust their hypotheses about correct their language use by actively producing the language and receiving feedback on it. Because students in CLIL classrooms are particularly encouraged to use the foreign language for communication, there is ample room for language output on the part of the students and for feedback by the teacher and peers.

With the natural approach, Krashen and Terrell (2000) provided another theoretical, methodologically oriented basis for the assumption of a positive CLIL effect on students' foreign-language competences. It draws on Krashen's (1985) monitor model and suggests that a foreign language can be effectively learned when the acquisition process occurs under natural conditions that resemble the setting in which the mother tongue was acquired. To ensure a natural language acquisition environment, Krashen and Terrell (2000) defined the following preconditions: First, to enable natural communication, the topics and classroom activities should be interesting and meaningful. A high level of authenticity is automatically found in CLIL classrooms because the content subjects' curricula predefine distinctive topics that need to be addressed, whereas the (foreign) language functions only as a tool for reaching this goal. Second, this focus on meaning

instead of on form implies that there should not be an undue involvement in grammatical features in the classroom. Moreover, speech errors are not corrected if the message can still be understood—an approach that is rooted in CLIL methodology, too. Third, students should not be forced to produce the second language but should be allowed to choose when to do so. In CLIL lessons, this corresponds to the opportunity to switch into the students' mother tongue if it seems necessary.

In addition to the above theoretical approaches, transferring Cummins' (1979) threshold hypothesis to the CLIL context suggests that students profit from CLIL not only in the foreign language but also in their content-subject competences. Cummins assumes that there are two thresholds that need to be crossed by learners to benefit from bilingualism: first, a lower threshold level (nativelike level in Language 1,⁵ above the low level in Language 2) to avoid negative cognitive effects, and second, a higher threshold level (high levels in Language 1 and Language 2) to obtain positive cognitive effects. This “additive bilingualism” can then lead to enhanced “metalinguistic, academic, and cognitive functioning” (Cummins, 2006, p. 182), which affects all scholastic domains, not only language competences. It can be argued that CLIL provides ideal conditions for students to cross this decisive second threshold (ZydatiB, 2012). First, high levels of L1 are assumed because it is their mother tongue. Second, CLIL students are likely to achieve high competences in the foreign language because they are typically exposed to a large amount of foreign language input in the CLIL setting (cf. Krashen, 1985) and have ample opportunities to interact with it (cf. Gass & Mackey, 2007). Thus, the consequent “linguistic and cognitive growth” (Cummins, 2006, p. 175) might positively influence the students' performance in the content subject.

To account for this cognitive growth, Adesope, Lavin, Thompson, and Ungerleider's (2010) meta-analysis of the cognitive correlates of bilingualism can be used. They found that bilingual children show higher metalinguistic and metacognitive awareness, more beneficial abstract and symbolic representations, more efficient attentional control, and better problem solving mechanisms than monolingual children. According to Adesope et al. (2010), these benefits are caused by the constant switching between the two languages that students perform in a bilingual setting. As a result, bilingual children are linguistically and cognitively more attentive, more creative, more flexible, and process information more deeply. However, it must be acknowledged that this is only partly applicable to CLIL students because they are not “true” bilinguals in the sense of growing up speaking two languages (Tej & William, 2008). Typically, CLIL students begin learning the foreign language in primary school and only in foreign-language lessons.

Previous Research on CLIL Effects on English and History Competences

Fehling (2008) investigated the English-language awareness of 139 secondary school students in Grades 7 and 8. In Grades 5 and 6, CLIL students underwent a preparatory course with additional

⁵ Language 1: mother tongue; Language 2: foreign language.

English lessons with bilingually taught lessons beginning in Grade 7. She found that the CLIL students already outperformed the non-CLIL group at the beginning of Grade 7 and also showed greater progress in language awareness over the course of Grade 8. Furthermore, students took a verbal cognitive skills test on which the CLIL students performed better than their non-CLIL peers. Yet, this fact was not controlled for when identifying their advantages in competence development. The extent to which the CLIL students' better English performance could be ascribed to their more favourable learning prerequisites or to CLIL practice was unknown.

By contrast, Nold et al. (2008) accounted for students' differing learning prerequisites by matching almost 2,000 CLIL and non-CLIL students by gender, mother tongue competence, basic cognitive skills, and socioeconomic status. They investigated students' English competences in Grade 9 and found that CLIL students showed higher initial levels of both listening comprehension and general English skills. However, improvement over Grade 9 was only larger for listening comprehension (DESI-Konsortium, 2006).

Rumlich (2014b) also investigated the English-skill development of 626 CLIL and non-CLIL secondary school students while controlling for gender, verbal cognitive abilities, interest, and academic self-concept in English. There were two comparison groups: non-CLIL students from the same schools that the CLIL students attended and from other schools that offered no CLIL programme. At the end of Grade 6, after a 2-year English preparatory course, CLIL students already performed better than the non-CLIL groups. Subsequently, after 2 years of bilingual teaching, CLIL students showed better competence development than their non-CLIL peers from the same schools. Yet, similarly to Nold et al. (2008), Rumlich found parallel competence development between CLIL students and the non-CLIL group from schools without a CLIL programme.

Regarding the bilingually taught content subject, there are hardly any studies that have investigated CLIL effects on student competence development (Bonnet, 2012; Coyle, Hood, & Marsh, 2010). In Jäppinen's (2005) study, 669 CLIL and non-CLIL students of different age levels (7, 10, and 13) completed a math and science test four times at intervals of half a school year. Jäppinen found that the CLIL and non-CLIL groups performed very similarly on the tests. Only the 10-year-old CLIL students performed significantly better than their non-CLIL peers at the third and fourth points of data collection.

These findings are in accordance with Badertscher and Bieri's (2009) results. They interviewed CLIL students in six secondary school classes to test students' content knowledge before and after monolingually and bilingually taught History lessons. This procedure was repeated after 1.5 years. The authors summarised that no differences were observed in the quality of knowledge produced regardless of whether the lessons were taught bilingually or monolingually. Yet, the findings are limited by the within-subjects design. The language of instruction was completely confounded with the respective topics. Nevertheless, the two studies indicated that there was no notable CLIL effect. The content-subject competences of both student groups seemed to develop

comparably. However, the above findings are limited by the fact that many CLIL programmes grant one additional lesson per week for the bilingually taught content subject compared with non-CLIL classrooms.

Finally, most cross-sectional studies have found that CLIL groups perform better in the content subject than non-CLIL students (e.g., Lamsfuß-Schenk, 2008; Madrid, 2011). Yet, because students' prior achievement was not considered, it was unknown whether performance differences had already existed from the onset; and as Dallinger and Jonkmann (2014) showed, such differences were likely. In general, no study has investigated CLIL effects on the content subject while controlling for confounding variables.

Confounders of the CLIL Effect at the Classroom Level

When isolating the effects of CLIL on student competences, it is important to control for possible differences in learning prerequisites (e.g., cognitive skills, motivation) and other aspects that are likely to differ between CLIL and non-CLIL classrooms (e.g., average classroom achievement) since they can be confounded with the CLIL effect. Ideally, this is done at the class level as CLIL always refers to entire classes and not only to individual students. Yet, no known studies have undertaken this. In the CLIL context, especially the consideration of teachers seems to be important. Because of a lack of appropriate CLIL-material, CLIL teachers need to invest far more time in preparing lessons than their non-CLIL colleagues (Viebrock, 2007; Zydatiņ, 2007). Therefore, Dalton-Puffer, Nikula, and Smit (2010) assumed that "CLIL-teachers are special in that they are willing to take on a considerable amount of extra work, which usually implies higher levels of motivation and pedagogical interest than teachers taken more generally" (p. 282). Thus, teacher enthusiasm (Kunter, Klusmann, Baumert, Richter, Voss, & Hachfeld, 2013) and self-efficacy (Holzberger, Philipp, & Kunter, 2013) should be controlled for when estimating the effects of CLIL on content learning. Furthermore, since the effects of teacher competences on student learning are mediated by the quality of the instruction that they provide (Kunter et al., 2013), instructional quality should be considered, too. This is true for the quality of instruction in not only the content subject but also regular English lessons because it can be argued that CLIL students attract especially motivated English teachers who in turn want to work with and make particular use of the students' high foreign-language competences.

Research Questions

Considering the current state of research, which is somewhat sparse and contradictory, there is a need for studies to investigate the competence development of CLIL and non-CLIL students in English and the content subject while controlling for confounding variables to isolate the CLIL effect as much as possible. This is COMBIH's research focus. Therefore, we conducted a longitudinal study with its first point of data collection (t1) at the beginning of bilingually taught History lessons in Grade 8

and the second (t2) at the end of Grade 8 after one school year of CLIL input. The COMBIH research questions were as follows: Does students' English competence development differ between CLIL and non-CLIL classrooms after 1 year of History lessons when considering prior achievement, learning prerequisites, and quality of English instruction? Does the competence development in History differ considering the same variables plus teacher characteristics? Concerning English competences, we expected that CLIL classrooms would show greater progress than non-CLIL groups even after differences in prior achievement, learning prerequisites, and quality of instruction had been controlled for. Effects should be especially strong regarding the receptive skill listening comprehension. Regarding History, we expected similar competence developments in CLIL and non-CLIL classrooms or a slightly negative CLIL effect due to the consideration of confounding variables.

Method

Sample

The sample comprised 1,806 eighth-graders from 37 German academic-track secondary schools in the state of Baden-Wuerttemberg (50.3% girls; $M_{age} = 13.5$, $SD_{age} = 0.51$); 68.8% of the students reported that at least one parent had qualified for a college education; 26.0% of the students reported having an immigration background (i.e., at least one parent was not born in Germany). The sample consisted of three groups. The first group comprised 703 students from 30 classes (28 schools) who took part in a CLIL programme. In Grades 5 and 6, they attended two additional English lessons per week. From Grade 7, at least one subject was taught bilingually throughout the school year. In Grade 7, this subject was Geography; in Grade 8, students took part in bilingual History lessons for the first time. They had already attended regular History lessons taught in their mother tongue in Grades 6 and 7. In Grade 8, the CLIL students' timetable was comprised of one extra History lesson per week compared with the monolingually taught students in the following two groups. The second group was comprised of 659 learners from 28 classes. They were from the same schools as the CLIL students but participated in a different branch. The 444 students (17 classes, nine schools) from the third group attended schools without a CLIL programme. The latter two groups participated in regular English lessons beginning in Grade 5 and in regular History lessons taught in their mother tongue beginning in Grade 6.

Seventeen of the CLIL-schools have a separate CLIL strand where CLIL students attend all subjects together with the same CLIL-classmates. In the other CLIL schools, CLIL students are pulled out of their home classrooms for the bilingually taught content lessons only. All other subjects – including English – are taught together with their non-CLIL classmates. Consequently, these schools needed to be excluded for the analyses on English learning because there is no clear classroom membership which is required for controlling for classroom-level factors (see Nold et al.,

2008, for the same approach). The reduced strand-sample comprised 1,281 students from 54 classrooms.

Procedure

The data were collected in autumn 2012 and summer 2013, at the beginning and the end of Grade 8. Beforehand, permission to conduct the study was obtained from the Ministry of Education and Cultural Affairs of Baden-Wuerttemberg. Student participation was voluntary, and written consent was obtained from their parents. Excluding breaks, data collection took 135 min and was conducted during regular classes by research assistants using a standardised manual. Students worked on the material by themselves in a paper-and-pencil format.

Instruments

English achievement. Students' overall English proficiency was assessed with a C-test (Klein-Braley, 1985). The present C-test consisted of 159 items (i.e., 159 gaps in every second word) that were distributed over seven subtexts with varying difficulty levels. To reduce test repetition effects, a booklet design was used. At t1, all students filled out Subtexts 1, 3, and 5, and half of the students, Subtexts 2 and 4. The other half received Subtexts 6 and 7 instead. At t2, students filled out the subtexts that they had not yet worked on and Subtexts 1, 3, and 5 again. For the psychometric analyses of the items and the IRT-based scaling of the test, a 2-parameter logistic model (Birnbaum model), which allows both the item difficulty and discrimination parameters to vary (Birnbaum, 1968) was chosen. Before estimating the item parameters, the items were examined with regard to Differential Item Functioning (DIF; variation in item parameters between different subpopulations) and Item Parameter Drift (IPD; variation in item parameters across time points) in order to guarantee fair and unbiased competence estimations. Gender (female vs. male) as well as language of instruction (monolingual vs. bilingual) were taken into consideration as DIF variables. The differences between group- and time-specific item parameters were tested for significance using the delta theorem. In addition, for each item, an estimation of the practical significance was conducted on the basis of graphic procedures in which the model-implied Item Characteristic Curves (ICCs) of the subgroups were compared with each other. If the ICCs differed clearly, this was interpreted as an indication of DIF, and the item was removed. This was true for 21 pretest items and 18 posttest items. The drift analysis was also conducted with regard to statistical and practical significance. The statistical test of the item parameters and the graphic inspection of the ICCs showed that, after excluding the DIF items, two of the remaining items deviated considerably between the pretest and posttest. For these items, a variation of the item parameters between pretest and posttest (i.e., partial measurement invariance) was allowed. With the remaining items and the resulting restrictions, the final measurement model was specified. Person parameters were estimated with the help of EAP estimators. To determine the reliability $Rel(EAP)$ of the tests, the variance of the person parameter

EAP was related to the latent ability dimension theta. The reliability was high at both time points with $\text{Rel(EAP)}_{t1} = .94$ and $\text{Rel(EAP)}_{t2} = .95$. The EAP estimators were t-standardised by the mean and standard deviation of the t1 scores.

Furthermore, English listening comprehension was assessed with items from Köller, Knigge, and Tesch (2010). The present test consisted of nine tasks with 43 open or closed items. At t1, Tasks 1 to 6 were employed. At t2, Tasks 4 and 5 were used again to serve as anchor items. In addition, we employed three new tasks that were more challenging to account for the students' learning progress. Again, the 2-parametric Birnbaum model was chosen (Birnbaum, 1968). Gender (female vs. male) and language of instruction (monolingual vs. bilingual) were taken into consideration as DIF variables, and DRIFT analyses were conducted. The ICCs of four pretest items differed clearly and were removed. None of the posttest items showed DIF. Regarding the drift analysis of the anchor items, the statistical test of the item parameters and the graphical inspection of the ICCs at t1 showed that none of them deviated considerably from those of the same items at t2. Consequently, all parameters of the anchor items were set equal between pretest and posttest. Then, the final measurement model was specified. EAP estimators were used for person parameters (t-standardisation; cf. above). At t1, the reliability was satisfactory, $\text{Rel(EAP)}_{t1} = .71$. At t2, the reliability was even higher, $\text{Rel(EAP)}_{t2} = .84$.

History achievement. Students' knowledge of history was assessed with a cloze test and multiple-choice questions totalling 36 items. They were developed according to the state's History curriculum. Students were asked for important dates, names, technical terms, and to identify correct historical relationships and the reasons for and consequences of historical events. The test was deployed at t1 as a prior knowledge test and at t2 to assess the students' learning progress. At t1, all students received the test in German. At t2, one third of the CLIL students received the test in German. The other CLIL students received either the cloze test or the multiple-choice questions in English. For the psychometric analyses of the items and the IRT-based scaling of the test, again, the 2-parametric Birnbaum model was chosen (Birnbaum, 1968). To avoid biasing the parameter estimation due to the variability in language, the English items were treated as separate items in the model, and their item parameters were allowed to differ from the German ones. DIF and DRIFT analyses were conducted. DIF was observed for three pretest items, two German posttest items, and two English posttest items. These items were removed. For the DRIFT analysis, the statistical test of the item parameters and a graphical inspection of the ICCs showed that the item parameters of eight items differed considerably between t1 and t2. For these items, the item parameters were allowed to vary between the pretest and posttest (i.e., partial measurement invariance). Then, the final measurement model was specified. EAP estimators were used for person parameters (t-standardisation; cf. above). At t1, the reliability was satisfactory, $\text{Rel(EAP)}_{t1} = .70$. At t2, the reliability was even higher, $\text{Rel(EAP)}_{t2} = .79$.

Cognitive abilities. At t1, students completed a figural and verbal cognitive skills test (KFT; Heller & Perleth, 2000) that asked them to detect matching figures and word pairs. Parallel test versions were employed to prevent students from copying their neighbours' results. The figural test comprised 25 items with reliabilities of $\alpha = .91$ and $\alpha = .90$ (Versions A and B, respectively). The verbal test comprised 20 items with $\alpha = .69$ and $\alpha = .72$, respectively.

Motivation. English and History motivation was measured based on the expectancy-value model of achievement motivation (Wigfield & Eccles, 2000) and applied at t1. Expectancy was measured by self-concept (four items; History $\alpha = .86$; English $\alpha = .84$; e.g., "I am good in History/English"). Value was assessed regarding both intrinsic value (five items; History: $\alpha = .82$; English: $\alpha = .72$; e.g., "History/English is one of my favourite subjects") and utility value (five items; History: $\alpha = .89$; English: $\alpha = .88$; e.g., "History/English is very useful for my future life"). Responses were given on a 4-point Likert-type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*).

Family background. The items for measuring the students' family background were adapted from large-scale studies such as PISA (Klieme, Artelt, Hartig, Jude, Köller, & Prenzel, 2010). On the basis of students' information about their parents' occupation and education, the International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, de Graaf, Treimann, & de Leeuw, 1992; HISEI for the higher ISEI of the two parents) was determined for each student, ranging between 11 and 88 (very low to very high socioeconomic status). Furthermore, information was collected on students' gender and immigration background (at least one parent was not born in Germany).

Quality of instruction. Quality of instruction in English and History was assessed as in the COACTIV study (Kunter et al., 2011) with respect to three basis dimensions: classroom management (seven items; History: $\alpha = .89$; English: $\alpha = .91$; e.g., "Our English/History lessons frequently start too late"), constructive teacher support (16 items; History: $\alpha = .87$; English: $\alpha = .86$; e.g., "Our English/History teacher does not think it is bad when we make mistakes"), and cognitive activation (seven items; History: $\alpha = .78$; English: $\alpha = .83$; e.g., "Our English/History teacher frequently makes us explain our thought processes"). Some items were adapted or added to ensure content-subject specificity. In addition, teacher enthusiasm (three items; History: $\alpha = .83$; English: $\alpha = .89$; e.g., "Our English/History teacher teaches enthusiastically") was assessed to indicate the teachers' overall motivation for teaching. Responses to all items were given on a 4-point Likert-type scale ranging from 1 (*completely disagree*) to 4 (*completely agree*).

Teacher characteristics. Teacher reports are available for 86% of the teachers. At t1, enthusiasm was measured with six items ($\alpha = .89$; e.g., "I really enjoy teaching"; Kunter, Frenzel, Nagy, Baumert, & Pekrun, 2011), self-efficacy with 10 items ($\alpha = .81$; e.g., "When I try really hard, I am able to reach even the most difficult students"; Schwarzer & Jerusalem, 1999), using a 4-point Likert-type scale (1 (*completely disagree*) to 4 (*completely agree*)).

Analyses

To examine possible differences in the competence development of CLIL and non-CLIL classrooms, two-level linear regression analyses were computed to take into account the hierarchical structure of the data in which individuals were nested within classes (Raudenbush & Bryk, 2002). Four models were estimated in Mplus 7.1 (Muthén & Muthén, 1998-2012) for each of the two English tests and the History test. Test performance at t2 was used as the dependent variable. In the first model (Model 1), test performance was predicted only by the two dummy variables at the class level (L2), indicating students' affiliation with a CLIL or one of the two non-CLIL classes. At the student level (L1), no predictors were included in Model 1. In Model 2, Model 1 was extended by prior achievement at L1 and L2. Model 3 expanded Model 2 in order to predict possible differences in competence development between CLIL and non-CLIL classrooms more comprehensively by controlling for learning prerequisites at both levels. In Model 4, Model 3 was extended by quality of instruction at L1 and L2 and teacher characteristics at L2. Before the analyses, all continuous variables were z-standardised. Afterwards, the continuous independent variables were group-mean centred, and all independent variables were aggregated at the classroom level by calculating their classroom means.

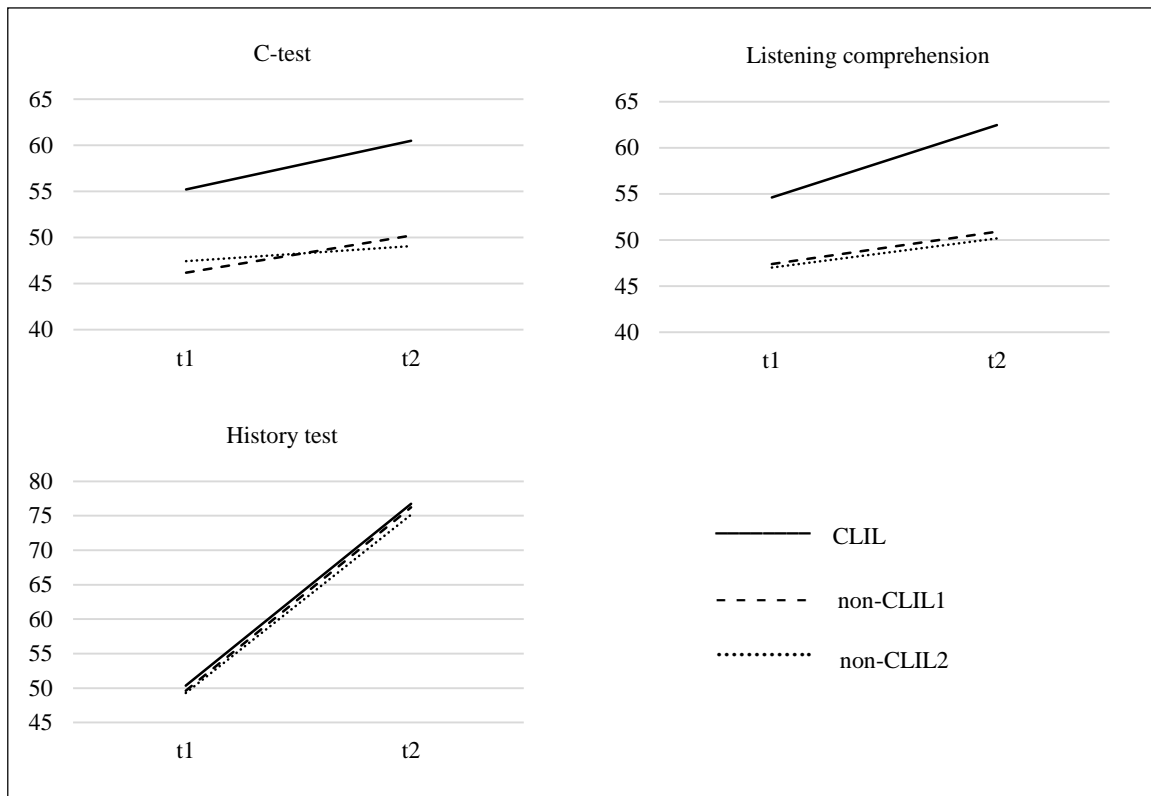
The percentage of missing values ranged from 0% (gender) to 21% (cognitive activation in English). The missing value estimator in Mplus was used to account for these missing values. Mplus applies a model-based approach, which builds on full information maximum likelihood (FIML) estimation.

Results

Descriptive Results

Figure 1 shows student competence development in English and History using the t-standardised group means. The students' performances improved on all three tests over the course of the school year. On the two English tests, the CLIL students performed considerably better at t1 and t2 than the non-CLIL students from CLIL schools (C-test: t1: $d = 0.90$; t2: $d = 0.82$; listening comprehension: t1: $d = 0.74$; t2: $d = 1.09$) and the non-CLIL students from schools without a CLIL programme (C-test: t1: $d = 0.78$; t2: $d = 0.81$; listening comprehension: t1: $d = 0.77$; t2: $d = 1.13$). Yet, only in listening comprehension CLIL students seem to have made a larger progress compared to non-CLIL students. Regarding History, the group means did not differ strongly between CLIL students (t1 = 50.5, t2 = 76.9), non-CLIL students from CLIL schools (t1 = 49.8, t2 = 76.0), and non-CLIL students from schools without a CLIL programme (t1 = 49.5, t2 = 75.2), thus suggesting comparable competence developments.

Fig. 1: Competence development on the achievement tests



Note: t1 and t2 = first and second points of data collection; CLIL = CLIL students; non-CLIL1 = non-CLIL students from CLIL schools; non-CLIL2 = non-CLIL students from schools without a CLIL programme.

Competence Development

Research Questions 1 and 2 addressed the competence development of CLIL and non-CLIL students in English and History. The results of the four two-level linear regression models will each be presented next. Table 1 (C-test & listening comprehension) and Table 2 (History test) display regression coefficients, standard errors, statistical significance tests, and the proportions of explained variance.

Student-level analyses. Before we report CLIL effects, we present the findings at the within-classroom student level (bottom halves of Tables 1 & 2). For all outcomes, prior achievement was the strongest predictor of future learning. Investigating all predictors simultaneously (Model 4), general abilities also incrementally predicted English achievement. The motivational variables showed differential relationships with the outcomes: self-concept was incrementally predictive of listening, intrinsic value of the C-test, and utility value for C-test and History knowledge. All other variables being equal, there was a significant gender difference favouring girls regarding C-test and boys regarding listening performance. Concerning instructional quality, students who reported higher teacher support showed better progress in the C-test and History. The student variables explained 55.2% of the total within-classroom variance of the t2-C-test, 29.3% for listening comprehension and 37.8% for History knowledge.

Between-classroom analyses. For the C-test (Table 1), Model 1 showed substantial achievement differences at t2 between CLIL and the non-CLIL classrooms (non-CLIL1/2: $B = -0.76/-0.84$; p 's $< .001$). Programme membership explained 54.2% of the total variance at the class level. Once differences in prior achievement at the beginning of Grade 8 were considered (Model 2), the advantage was substantially reduced and became statistically insignificant. Prior achievement explained an additional 25.8% of the classroom variance. The addition of further control variables in Models 3 (general abilities, motivation, demographics) and 4 (quality of instruction) explained an extra 8% of classroom variance but left the CLIL effect widely unaffected. In the full model, achievement differences between CLIL and non-CLIL classrooms were either insignificant (non-CLIL1: $B = -0.14$; $p = .176$) or only marginally significant (non-CLIL2: $B = -0.25$; $p = .073$).

For listening comprehension (Table 1), CLIL classrooms had higher average achievement than the non-CLIL classrooms (non-CLIL1/2: $B = -1.06/-1.13$; p 's $< .001$). Group membership explained 80.9% of the total classroom variance. In Model 2, this CLIL effect was still statistically significant, but greatly diminished to about 56% of its original size after prior achievement was controlled for. Prior achievement explained an additional 9.8% of the classroom variance. The consideration of further control variables in Models 3 and 4 explained an extra 6.9% of classroom variance. The CLIL effect only changed slightly but still indicated substantial advantages for CLIL classrooms. In the full model, classroom differences in listening were significantly predicted by average prior achievement, verbal abilities, intrinsic and utility value, and classroom management. That was also the case for classroom differences in the C-test (except classroom management).

Regarding History (Table 2), Model 1 indicated comparable levels of historical knowledge for the three student groups with small but statistically insignificant disadvantages for the non-CLIL classrooms (non-CLIL1/2: $B = -0.07/-0.19$; $p = .600/.118$). Accordingly, programme type only explained 2.8% of the total classroom variance. Controlling for prior achievement (Model 2) further reduced these slight disadvantages. It also explained a substantial amount of additionally 34.5% of the total classroom variance. Considering further control variables in Models 3 and 4, explained an additional 39% of the total classroom variance but left the CLIL effect widely unaffected: it remained statistically insignificant (non-CLIL1/2: $B = 0.12/-0.02$; $p = .321/.864$). In the full model, classroom differences in historical knowledge were significantly predicted by average prior achievement in the Grade 8 test, proportion of student with immigration background, classroom management, and cognitive activation.

Tab. 1: Development of C-test and listening comprehension competences

| | C-test | | | | | | | | Listening comprehension | | | | | | | |
|---|----------|-----------|----------|-----------|----------|-----------|----------|-----------|-------------------------|-----------|----------|-----------|----------|-----------|----------|-----------|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
| | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> |
| Class level | | | | | | | | | | | | | | | | |
| Reference group | 0.52 | ***0.08 | 0.08 | 0.06 | 0.27 | *0.12 | 0.21 | 0.12 | 0.70 | ***0.07 | 0.41 | ***0.11 | 0.65 | ***0.12 | 0.60 | ***0.11 |
| Non-CLIL1 | -0.76 | ***0.11 | 0.00 | 0.09 | -0.15 | 0.10 | -0.14 | 0.10 | -1.06 | ***0.10 | -0.60 | ***0.16 | -0.62 | ***0.12 | -0.61 | ***0.11 |
| Non-CLIL2 | -0.84 | ***0.14 | -0.19 | 0.12 | -0.25 | 0.15 | -0.25 | 0.14 | -1.13 | ***0.10 | -0.65 | ***0.15 | -0.67 | ***0.11 | -0.66 | ***0.11 |
| Prior achievement | | | 0.81 | ***0.09 | 0.48 | ***0.10 | 0.52 | ***0.11 | | | 0.65 | ***0.15 | 0.52 | ***0.14 | 0.54 | ***0.13 |
| Figural skills | | | | | 0.14 | 0.18 | 0.08 | 0.19 | | | | | 0.07 | 0.17 | 0.02 | 0.16 |
| Verbal skills | | | | | 0.32 | *0.15 | 0.34 | *0.17 | | | | | 0.26 | 0.17 | 0.32 | *0.16 |
| Self-concept | | | | | 0.17 | 0.19 | 0.17 | 0.19 | | | | | 0.02 | 0.11 | 0.02 | 0.10 |
| Intrinsic value | | | | | -0.29 | *0.15 | -0.36 | *0.14 | | | | | -0.30 | *0.11 | -0.33 | *0.11 |
| Utility value | | | | | 0.30 | *0.13 | 0.32 | *0.13 | | | | | 0.34 | *0.12 | 0.36 | *0.11 |
| Gender (1=female) | | | | | -0.49 | 0.25 | -0.37 | 0.26 | | | | | -0.29 | 0.16 | -0.19 | 0.16 |
| Immigration (1=yes) | | | | | 0.16 | 0.25 | 0.11 | 0.23 | | | | | -0.09 | 0.22 | -0.15 | 0.20 |
| HISEI | | | | | 0.09 | 0.12 | 0.06 | 0.11 | | | | | 0.18 | 0.10 | 0.13 | 0.09 |
| Classroom management | | | | | | | 0.10 | 0.07 | | | | | | | 0.10 | *0.04 |
| Teacher support | | | | | | | -0.03 | 0.09 | | | | | | | -0.07 | 0.05 |
| Cognitive activation | | | | | | | 0.04 | 0.11 | | | | | | | 0.05 | 0.11 |
| Student level | | | | | | | | | | | | | | | | |
| Prior achievement | | | 0.71 | ***0.04 | 0.61 | ***0.04 | 0.60 | ***0.04 | | | 0.42 | ***0.03 | 0.31 | ***0.03 | 0.31 | ***0.03 |
| Figural skills | | | | | 0.06 | *0.03 | 0.06 | *0.03 | | | | | 0.07 | **0.02 | 0.07 | **0.02 |
| Verbal skills | | | | | 0.13 | ***0.03 | 0.12 | ***0.03 | | | | | 0.14 | ***0.03 | 0.14 | ***0.03 |
| Self-concept | | | | | 0.05 | 0.03 | 0.05 | 0.03 | | | | | 0.18 | ***0.03 | 0.18 | ***0.03 |
| Intrinsic value | | | | | -0.04 | 0.02 | -0.05 | *0.03 | | | | | -0.05 | 0.03 | -0.05 | 0.04 |
| Utility value | | | | | 0.09 | **0.03 | 0.09 | **0.03 | | | | | 0.04 | 0.03 | 0.04 | 0.03 |
| Gender (1=female) | | | | | 0.14 | ***0.04 | 0.14 | ***0.04 | | | | | -0.09 | *0.04 | -0.09 | *0.04 |
| Immigration (1=yes) | | | | | -0.03 | 0.05 | -0.01 | 0.05 | | | | | -0.10 | 0.06 | -0.09 | 0.06 |
| HISEI | | | | | -0.01 | 0.02 | -0.02 | 0.02 | | | | | 0.02 | 0.02 | 0.02 | 0.02 |
| Classroom management | | | | | | | 0.01 | 0.04 | | | | | | | 0.01 | 0.03 |
| Teacher support | | | | | | | 0.13 | **0.05 | | | | | | | 0.03 | 0.04 |
| Cognitive activation | | | | | | | -0.02 | 0.02 | | | | | | | -0.01 | 0.03 |
| R ² class level ^a | 54.2 | | 80.0 | | 86.7 | | 88.0 | | 80.9 | | 90.7 | | 96.5 | | 97.6 | |
| R ² student level | | | 49.4 | | 54.0 | | 55.2 | | | | 20.8 | | 29.3 | | 29.3 | |

Note. *B* = regression coefficient, *SE* = standard error; non-CLIL1 = non-CLIL students from CLIL schools, non-CLIL2 = non-CLIL students from schools without CLIL programme; HISEI = International Socio-Economic Index of Occupational Status.

^aIntraclass correlation coefficient C-test = 0.27; listening comprehension = 0.34.

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

Tab. 2: Development of historical knowledge

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
|---|----------|-----------|----------|-----------|----------|-----------|----------|-----------|
| | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> | <i>B</i> | <i>SE</i> |
| Class level | | | | | | | | |
| Reference group | 0.09 | 0.09 | 0.02 | 0.08 | 0.14 | 0.25 | 0.09 | 0.19 |
| Non-CLIL1 | -0.07 | 0.13 | 0.05 | 0.12 | 0.16 | 0.15 | 0.12 | 0.12 |
| Non-CLIL2 | -0.19 | 0.12 | -0.11 | 0.10 | -0.08 | 0.12 | -0.02 | 0.11 |
| Prior achievement | | | 0.41 | **0.14 | 0.23 | 0.14 | 0.22 | *0.11 |
| History test Grade 7 | | | 0.32 | *0.13 | 0.16 | 0.10 | 0.06 | 0.07 |
| Figural skills | | | | | 0.29 | 0.21 | 0.28 | 0.20 |
| Verbal skills | | | | | 0.27 | 0.27 | 0.33 | 0.24 |
| Self-concept | | | | | -0.03 | 0.32 | 0.10 | 0.16 |
| Intrinsic value | | | | | -0.47 | 0.37 | 0.12 | 0.23 |
| Utility value | | | | | 0.18 | 0.16 | -0.27 | 0.22 |
| Gender (1=female) | | | | | -0.10 | 0.23 | 0.14 | 0.24 |
| Immigration (1=yes) | | | | | 0.12 | 0.24 | -0.63 | *0.31 |
| HISEI | | | | | 0.16 | 0.26 | 0.01 | 0.15 |
| Classroom management | | | | | | | 0.18 | **0.07 |
| Teacher support | | | | | | | -0.05 | 0.10 |
| Cognitive activation | | | | | | | 0.41 | ***0.11 |
| Teacher enthusiasm | | | | | | | -0.03 | 0.04 |
| Teacher self-efficacy | | | | | | | -0.02 | 0.04 |
| Student level | | | | | | | | |
| Prior achievement | | | 0.31 | ***0.03 | 0.28 | ***0.03 | 0.28 | ***0.03 |
| History test Grade 7 | | | 0.36 | ***0.03 | 0.31 | ***0.03 | 0.31 | ***0.03 |
| Figural skills | | | | | 0.01 | 0.03 | 0.00 | 0.02 |
| Verbal skills | | | | | 0.03 | 0.03 | 0.03 | 0.03 |
| Self-concept | | | | | 0.03 | 0.03 | 0.02 | 0.03 |
| Intrinsic value | | | | | 0.05 | 0.03 | 0.04 | 0.03 |
| Utility value | | | | | 0.06 | *0.03 | 0.06 | *0.03 |
| Gender (1=female) | | | | | -0.06 | 0.05 | -0.06 | 0.05 |
| Immigration (1=yes) | | | | | -0.03 | 0.05 | -0.02 | 0.05 |
| HISEI | | | | | 0.02 | 0.02 | 0.01 | 0.02 |
| Classroom management | | | | | | | 0.04 | 0.02 |
| Teacher support | | | | | | | 0.09 | **0.03 |
| Cognitive activation | | | | | | | 0.04 | 0.02 |
| R ² class level ^a | 2.8 | | 37.3 | | 54.0 | | 76.3 | |
| R ² student level | | | 35.1 | | 36.6 | | 37.8 | |

Note. *B* = regression coefficient, *SE* = standard error; non-CLIL1 = non-CLIL students from CLIL schools, non-CLIL2 = non-CLIL students from schools without CLIL programme; HISEI = International Socio-Economic Index of Occupational Status.

^aIntraclass correlation coefficient = 0.19.

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

Discussion

The aim of this article was to investigate the effect of CLIL on students' English and History competences. According to our expectations, competence development in listening comprehension was substantially greater in CLIL than in non-CLIL classrooms over the course of Grade 8 even though confounding variables (prior achievement, learning prerequisites, quality of instruction) were controlled for. For general English skills, we only found marginally to not significant advantages for the CLIL-programme (Research Question 1). Regarding History, the results displayed, as expected, comparable competence developments between CLIL and non-CLIL classrooms (Research Question 2).

COMBIH replicated the results of several previous studies (Fehling, 2008; Nold et al., 2008; Rumlich, 2014b), which found that CLIL had a positive effect on students' English skills – mainly on receptive skills such as listening comprehension. This finding is not surprising considering the CLIL students enhanced exposure to English throughout the school year. Yet, this CLIL effect had a tendency to be overestimated in previous studies due to the failure to consider confounding variables (Costa & D'Angelo, 2011; Pérez-Cañado, 2012). Therefore, COMBIH aimed to further validate previous findings by controlling for prior achievement, learning prerequisites, and quality of instruction. We found that CLIL classrooms still showed substantially better competence developments in listening comprehension than the non-CLIL groups. These results confirm a positive CLIL effect and allow the transfer of theoretical approaches of second-language acquisition to the CLIL context. Accordingly, a sufficient amount of understandable foreign-language input (Krashen, 1985) and the acquisition of a foreign language in an environment that resembles the setting in which the mother tongue was acquired (Krashen & Terrell, 2000)—as implemented in the CLIL setting—lead to enhanced foreign-language competences of CLIL students. Regarding competence development on the C-test, however, the consideration of confounding variables reduced the CLIL effect strongly and it, thus, became statistically only marginally significant or nonsignificant. One reason for this nonsignificant CLIL effect might be that C-tests require active language production instead of reception. To investigate this hypothesis further, future studies should also measure reading, speaking and writing skills, which was not possible in this large-scale study.

Regarding History, the present results confirm Jäppinen's (2005) and Badertscher and Bieri's (2009) findings of comparable competence developments between CLIL and non-CLIL groups. Thus, thanks to COMBIH's longitudinal design, we were able to confirm the few results on this matter as opposed to cross-sectional studies, which tend to overestimate the CLIL effect due to a failure to consider prior achievement. Moreover, the present study also controlled for several confounding variables for the first time. This did not lead to different results. Consequently, CLIL does not seem to cause either positive or negative effects in History. This result contradicts the assumptions of Cummins' (1979) threshold hypothesis and Adesope et al.'s (2010) findings, which

claim that a bilingual instructional setting provides the students with cognitive benefits, which in turn might result in higher content-subject competences. One reason why we did not find such advantages might be that CLIL students are not “true” bilinguals (cf. Tej & William, 2008) because they learn English at school only. Thus, it seems that the benefits of cognitive correlates of bilingualism are not applicable to the “artificial” bilingual setting of CLIL. Future studies are needed to confirm these findings.

Yet, the lack of a negative CLIL effect can also be positively evaluated. CLIL students demonstrated the same skills even though they had to master linguistic challenges in addition to the regular demands that the acquisition of new content brings along. This interpretation, however, must be treated with caution given the extra History lesson per week that is included in most CLIL classes' timetables. Consequently, a negative CLIL effect might have been found if both student groups had been given the same number of History lessons. The slight but statistically nonsignificant advantages in favour of the non-CLIL groups that were found in present analyses after the confounding variables had been considered supported this assumption. Future studies would have to examine CLIL effects in samples that are given the same number of content-subject lessons. Nevertheless, potential performance disadvantages of CLIL classrooms could be relativised through test language. Saalbach, Eckstein, Andri, Hobi, and Grabner (2013) showed that instruction and assessment in two different languages led to cognitive costs. This might have been the case in COMBIH. Therefore, more research is needed in which the languages of instruction and assessment are consistent.

In the present study, only a small part of History competences, namely content knowledge, could be assessed. It is important that future research conducts similar analyses that focus on other skills, such as the ability to consider and evaluate a historical event from multiple perspectives.

Conclusion

Even though confounding variables could not be considered exhaustively and no cause-effect statements could be made, COMBIH aimed to isolate the CLIL effect on student competences as much as possible by applying a longitudinal design and by controlling for various confounding variables. Students' listening comprehension skills seem to benefit from CLIL, CLIL seems to lead to minor and therefore negligible advantages regarding general English skills, but CLIL does not seem to be detrimental to their content-subject achievement. Thus, there is a positive CLIL effect but only with regard to receptive English skills. These results might provide a foundation for future decisions about the expansion of CLIL programmes. Further research should be conducted in other countries and should focus on different age groups and content subjects to broaden reliable insights into the effects of CLIL on student competences.

References

- Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*, 207–245. doi:10.3102/0034654310368803
- Badertscher, H., & Bieri, T. (2009). *Wissenserwerb im Content and Language Integrated Learning [Acquisition of knowledge in Content and Language Integrated Learning]*. Bern: Haupt.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Reading, MA: Addison-Wesley.
- Bonnet, A. (2012). Towards an evidence base for CLIL: How to integrate qualitative and quantitative as well as process, product and participant perspectives in CLIL research. *International CLIL Research Journal, 4*, 66–78. Retrieved from <http://www.icrj.eu/14/article7.html> on 25th November 2014
- Breidbach, S., & Viebrock, B. (2012). CLIL in Germany: Results from recent research in a contested field of education. *International CLIL Research Journal, 4*, 5–16. Retrieved from <http://www.icrj.eu/14/article1.html> on 25th November 2014
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System, 39*, 523–532. doi:10.1016/j.system.2011.08.002
- Costa, F., & D'Angelo, L. (2011). CLIL: A suit for all seasons. *Latin American Journal of Content & Language, 4*, 1–13. doi:10.5294/lacil.2011.4.1.1
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge UP.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222–251. doi:10.3102/00346543049002222
- Cummins, J. (2006). *Language, power and pedagogy*. Clevedon: Multilingual Matters.
- Dallinger, S., & Jonkmann, K. (2014). Lernausgangslagen von Schülerinnen und Schülern zu Beginn des bilingualen Geschichtsunterrichts [Students' learning prerequisites at the beginning of CLIL History lessons]. In C. Fäcke, M. Rost-Roth & E. Thaler (Eds.), *Sprachenausbildung – Sprachen bilden aus – Bildung aus Sprachen* (pp. 173–183). Baltmannsweiler: Schneider Hohengehren.
- Dalton-Puffer, C., Nikula, T., & Smit, U. (2010). Language use and language learning in CLIL. Current findings and contentious issues. In C. Dalton-Puffer, T. Nikula & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 279–291). Amsterdam: Benjamins.
- DESI-Konsortium (Ed.). (2006). *Unterricht und Kompetenzerwerb in Deutsch und Englisch (DESI) [Lessons and acquisition of competences in German and English (DESI)]*. Frankfurt/M.:

- DIPF. Retrieved from <http://www.dipf.de/de/projekte/pdf/biqua/desi-zentrale-befunde> on 27th November 2014
- Dumont, H., Neumann, M., Maaz, K., & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen [The effect of student body composition on academic achievement]. *Psychologie in Erziehung und Unterricht*, 3, 163–183. doi:0.2378/peu2013.art14d
- Eurydice (Ed.). (2006). *Content and Language Integrated Learning (CLIL) at school in Europe*. Retrieved from <http://bookshop.europa.eu/en/content-and-language-integrated-learning-clil-at-school-in-europe-pbNCX106001> on 19th September 2014
- Fehling, S. (2008). *Language Awareness und bilingualer Unterricht [Language awareness and CLIL]*. Frankfurt/M.: Lang.
- Ganzeboom, H., de Graaf, P., Treimann, D., & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1–56. doi:10.1016/0049-089x(92)90017-b
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175–199). Mahwah, NJ: Erlbaum.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Heller, K., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. – 12. Klassen. Revision [Cognitive abilities test for 4th – 12th-graders. Revision]*. Göttingen: Hogrefe.
- Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality. A longitudinal analysis. *Journal of Educational Psychology*, 105, 774–786. doi:10.1037/a0032198
- Jäppinen, A.-K. (2005). Thinking and content learning of mathematics and science as cognitional development in Content and Language Integrated Learning (CLIL). Teaching through a foreign language in Finland. *Language and Education*, 19, 148–169. doi:10.1080/09500780508668671
- Klein-Braley, C. (1985). A close-up on the C-test. A study in the construct validation of authentic tests. *Language Testing*, 2, 76–104. doi: 10.1177/026553228500200108
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., & Prenzel, M. (Eds.). (2010). *PISA 2009*. Münster: Waxmann.
- Köller, O., Knigge, M., & Tesch, B. (Eds.). (2010). *Sprachliche Kompetenzen im Ländervergleich [Comparison of linguistic competences in German states]*. Münster: Waxmann.
- Köller, O., Leucht, M., & Pant, H. (2012). Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I. [Effects of bilingual instruction on English achievement in lower secondary schools]. *Unterrichtswissenschaft*, 4, 334–350.

- Krashen, S. D. (1985). *The input hypothesis. Issues and implications*. London: Longman.
- Krashen, S. D., & Terrell, T. D. (2000). *The natural approach. Language acquisition in the classroom*. Harlow: Pearson.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften [Teachers' Professional Competence]*. Münster: Waxmann.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology, 105*, 805–820. doi:10.1037/a0032583
- Lamsfuß-Schenk, S. (2008). *Fremdverstehen im bilingualen Geschichtsunterricht [Understanding otherness in CLIL in History]*. Frankfurt/M.: Lang.
- Madrid, D. (2011). Monolingual and bilingual students' competence in social sciences. In D. Madrid & S. Hughes (Eds.), *Studies in bilingual education* (pp. 195–222). Bern: Lang.
- Muthén, L., & Muthén, B. (1998–2012). *Mplus user's guide*. Los Angeles: Muthén & Muthén.
- Nold, G., Hartig, J., Hinz, S., & Rossa, H. (2008). Klassen mit bilinguaem Sachfachunterricht. Englisch als Arbeitssprache [CLIL-classrooms. English as language of communication]. In DESI-Konsortium (Ed.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch* (pp. 451–457). Weinheim: Beltz.
- Pérez-Cañado, M. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism, 15*, 315–341. doi:10.1080/13670050.2011.630064
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models*. Thousand Oaks: Sage.
- Rumlich, D. (2014a). Prospective CLIL and non-CLIL students' interest in English (classes). In R. Breeze, C. Pasamar, C. Saíz, & C. Sala (Eds.), *Integration of theory and practice in CLIL* (pp. 75–95). Amsterdam: Rodopi.
- Rumlich, D. (2014b, March): Selektion und Präparation als Ursache für die hohen Sprachkompetenzen bilingual unterrichteter Lerner? [Selection and preparation as causes of CLIL students' high English competences?]. Paper presented at the meeting of the Gesellschaft für empirische Bildungsforschung (GEBF) in Kiel, Germany.
- Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction, 26*, 36–44. doi:10.1016/j.learninstruc.2013.01.002
- Tej, B. K., & William, R. C. (2008). *The handbook of bilingualism*. Oxford: Wiley-Blackwell.
- Viebrock, B. (2007). *Bilingualer Erdkundeunterricht. Subjektive didaktische Theorien von Lehrerinnen und Lehrern [CLIL in Geography. Teachers' subjective theories]*. Frankfurt/M.: Lang.
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. doi:10.1006/ceps.1999.1015

Zydati, W. (2007). *Deutsch-Englische Zge in Berlin (DEZIBEL)[German-English profiles in Berlin (DEZIBEL)]*. Frankfurt/M.: Lang.

Zydati, W. (2012). Linguistic thresholds in the CLIL classroom? The threshold hypothesis revisited. *International CLIL Research Journal*, 4, 17–28. Retrieved from <http://www.icrj.eu/14/article2.html> on 27th November 2014

Merkmale erfolgreichen bilingualen Sachfachunterrichts

– Die Rolle der Sprachen im deutsch-englischen
Geschichtsunterricht

Sara Dallinger

Abteilung Englisch, Pädagogische Hochschule Ludwigsburg

Kathrin Jonkmann

Abteilung Empirische Bildungsforschung und Pädagogische Psychologie, Universität Tübingen
Lehrgebiet Bildungspsychologie, FernUniversität Hagen

Jan Hollm

Abteilung Englisch, Pädagogische Hochschule Ludwigsburg

Christiane Fiege

Arbeitsbereich Schulpsychologie, Universität Tübingen

Dallinger, Sara; Jonkmann, Kathrin; Hollm, Jan & Fiege, Christiane (2015). *Merkmale erfolgreichen bilingualen Sachfachunterrichts – Die Rolle der Sprachen im deutsch-englischen Geschichtsunterricht*. Zur Veröffentlichung eingereichtes Manuskript.
(Die Zitation folgt den Vorgaben des Journals.)

Abstract

Trotz vieler Untersuchungen, die in den letzten Jahren zu bilinguaalem Sachfachunterricht durchgeführt wurden, sind kaum Studien zur Wirksamkeit spezifischer didaktisch-methodischer Unterrichtsaspekte bekannt. Dies trifft auch auf die Rolle der Sprachen zu, was das Spezifikum dieser Unterrichtsform darstellt. Deshalb untersuchte die vorliegende Studie die Effekte des Einsatzes von Mutter- und Fremdsprache auf die Leistungsentwicklung in Englisch und im Sachfach Geschichte von 30 bilingual unterrichteten 8. Klassen ($N = 703$) mithilfe von Lehrerangaben zur Häufigkeit des Englischesinsatzes durch Lehrkräfte und Lerner, zweckgebundener Verwendung des Deutschen sowie Sprachwahl bei Einführung von Fachbegriffen. Die Ergebnisse zeigten, dass sich sowohl vermehrter Englischinput durch die Lehrkräfte als auch der Miteinbezug des Deutschen günstig auf die Leistungsentwicklung in Englisch und Geschichte auswirkten. Ein großer englischer Redeanteil seitens der Lerner schien zu schlechteren Leistungsentwicklungen zu führen. Diese Befunde leisten einen Beitrag zur bislang noch weitestgehend ausstehenden Entwicklung einer empirisch basierten Methodik des bilingualen Unterrichts.

Merkmale erfolgreichen bilingualen Sachfachunterrichts – Die Rolle der Sprachen im deutsch-englischen Geschichtsunterricht

Einleitung

Der Terminus „bilingual“ verweist in der Regel auf die Präsenz zweier Sprachen in verschiedensten Kontexten. Im bilingualen Sachfachunterricht – bei dem Sachfächer, wie beispielsweise Geschichte oder Geographie, für gewöhnlich mithilfe zweier Sprachen unterrichtet werden – sind dies Mutter- und Fremdsprache. Doch in welchem Umfang werden die beiden Sprachen eingesetzt? Und wie wirkt sich der Einsatz der Sprachen auf die Leistungen der Schülerinnen und Schüler aus? Trotz umfassender Forschung zu bilingualem Unterricht scheinen diese Fragen bis dato weitgehend unbeantwortet. Denn viele Studien beschäftigten sich bisher mit den Kompetenzen von bilingual unterrichteten Lernern in Englisch (vgl. Dallinger et al. 2014; Köller et al. 2012; Nold et al. 2008; Zydati 2007) und in verschiedenen Sachfächern (vgl. Badertscher und Bieri 2009; Dallinger et al. 2014; Zydati 2007). Was bisher jedoch kaum untersucht wurde, sind Faktoren, die einen guten bilingualen Unterricht kennzeichnen und zu diesen Leistungen führen. Hierbei stellt die Rolle der Sprachen einen zentralen Aspekt dar, da es das Spezifikum dieser Unterrichtsform beschreibt. Trotz der Aktualität der Debatte (vgl. Diehr 2012) wurde die Sprachverwendung im bilingualen Unterricht und deren Auswirkungen bisher noch nicht zufriedenstellend erforscht: „Eine systematische und empirisch fundierte Methodik des BU [Bilingualen Unterrichts], die den Einbezug beider Sprachen erläutert und konkretisiert, bleibt ein Desiderat“ (Diehr 2012, S. 29).

Die vorliegende Studie „Competences and Motivation in Bilingual Instruction in History“ (COMBIH) möchte einen Beitrag zu diesem Thema leisten. Sie analysierte mittels Mehrebenenanalysen den Leistungszuwachs in Englisch und Geschichte in 30 bilingualen Gymnasialklassen im Verlauf eines Schuljahres in Abhängigkeit von methodisch-didaktischen Aspekten des Spracheinsatzes. Dabei machten die Lehrkräfte Angaben zur prozentualen Häufigkeit des Einsatzes von Englisch versus Deutsch seitens der Lehrkraft und der Lerner, zur zweckorientierten Wahl der Sprache in verschiedenen Unterrichtssituationen sowie zur Sprachwahl bei Einführung neuer Fachbegriffe.

Theoretischer Hintergrund

Aktueller Diskurs zur Zweisprachigkeit im bilingualen Unterricht

Obwohl der Terminus „bilingual“ auf Zweisprachigkeit im Unterricht hindeutet, scheint bilingualer Sachfachunterricht häufig als eine einsprachig in der Fremdsprache gehaltene Unterrichtsform verstanden zu werden. Dies bildet sich in den Bildungsplänen der Bundesländer Bremen, Hessen, Niedersachsen und Rheinland-Pfalz ab (vgl. Deutscher Bildungsserver 1996-2014) sowie in der Definition dieser Unterrichtsform durch das europäische Informationsnetzwerk Eurydice, in der die

Muttersprache unerwähnt blieb (Eurydice 2006, S. 8). Darüber hinaus scheint auch in der Schulpraxis bilingualer Unterricht überwiegend einsprachig in der Fremdsprache abzulaufen (Böing und Palmen 2012, S. 73; Diehr 2012, S. 18), was ebenfalls Befürworter im fachdidaktischen Diskurs findet (vgl. Heine 2010). Neben dieser einsprachig fremdsprachigen Auslegung von bilinguaalem Unterricht besteht eine weitere Position, die den Einsatz der Muttersprache als „Verständnis-Notnagel“ sowie für spezifisches Fachvokabular toleriert. Beispielsweise räumten Bonnet (2007, S. 130) und Theis (2010, S. 52) ein, dass sich die Muttersprache im Zusammenhang mit Fachbegriffen als hilfreich erweisen könne. Diese Haltung wird auch von den Bildungsplänen in Hamburg, Sachsen und Thüringen vertreten (vgl. Deutscher Bildungsserver 1996-2014). Ähnliche Konkretisierungen erfolgten in Informationsbroschüren des Kultusministeriums Baden-Württemberg, die erläuterten, dass die Erreichung kognitiv anspruchsvoller Lernziele „einen Rückgriff auf die deutsche Sprache erforderlich machen“ könne (KM 2008, S. 6; vgl. KM 2006). Eine dritte Position spricht sich für die reflektierte und systematische Mitbenutzung der Muttersprache aus. Durch ihre Verwendung werden Vorteile im Lernprozess der Schülerinnen und Schüler durch systematische Unterstützung der Bereiche Verständnis, Ausdruck und Fachterminologie erwartet (Diehr 2012, S. 25; Gierlinger, Hametner und Spann 2007, S. 37; Otten und Wildhage 2007, S. 19). Der Einbezug der Muttersprache wird daher als unverzichtbar bewertet (Heimes 2013, S. 349). Dieses insgesamt eher unklare Bild bezüglich der Rolle der Sprachen im bilingualen Unterricht ergibt sich möglicherweise aus der Tatsache, dass bisher keine einheitliche methodisch-didaktische Grundlage für diese Unterrichtsform geschaffen wurde (Thürmann 2013, S. 230; Breidbach 2010, S. 174).

Der Einsatz von Mutter- und Fremdsprache im Hinblick auf Schülerleistung: Das Prinzip der aufgeklärten Einsprachigkeit und Zweitspracherwerbstheorien

Das Prinzip der aufgeklärten Einsprachigkeit von Butzkamm (1973; bzw. Prinzip der funktionalen Fremdsprachigkeit, Butzkamm 2002) besagt, dass sich der in Abschnitt 2.1 zuletzt vorgestellte Ansatz, also der regelmäßige und reflektierte Einsatz der Muttersprache, günstig auf den Fremdsprachenerwerb der Lerner auswirken könnte. Butzkamm (2002) argumentierte, dass jeder Mensch bereits muttersprachliches Vorwissen erworben hat (z. B. semantisch-situatives und syntaktisch-formales Vorwissen, Lese- und Schreibfähigkeit) und bei Kontakt mit einer Fremdsprache automatisch Verknüpfungen dazu vornehmen würde. Dies erleichtere den Fremdspracherwerbsprozess, weshalb die muttersprachlichen „Vorleistungen“ aktiv genutzt und nicht unterdrückt werden sollten. Im Kontext des bilingualen Unterrichts, bei dem die Verbesserung der Fremdsprachenkenntnisse der Lerner eine zentrale Rolle spielt, kann dieses Prinzip umgesetzt werden, da die Muttersprache in dieser Unterrichtsform mitverwandt werden sollte (vgl. Abschnitt 2.1). Auch Butzkamm (2002, S. 150) bezeichnete den bilingualen Unterricht als „Idealfall“ für die Anwendung seines Konzepts aufgrund der dort vorherrschenden, besonders bedeutungsvollen und authentischen Kommunikation, die dem Fremdsprachenerwerbsprozess zusätzlich zugutekäme. Folglich scheint es, als ob Lerner durch den Miteinbezug der Muttersprache ihre

Fremdsprachenkenntnisse im bilingualen Unterricht besonders effektiv verbessern könnten. Dasselbe dürfte auf die Leistungen im Sachfach zutreffen. Denn es kann angenommen werden, dass das verbesserte Verstehen der Sprache im Unterricht sowie das Behandeln authentischer Themen zu einem gesteigerten Verständnis und größerer Akzeptanz der Kommunikationsinhalte führen.

Im Gegensatz zur Muttersprache ist der Einsatz der Fremdsprache im bilingualen Unterricht unumstritten. Unklarheiten bestehen jedoch darüber, in welchem Umfang und auf welche Weise die Fremdsprache verwendet werden soll (Diehr 2012, S. 29). Verschiedene theoretische Ansätze zum Zweitspracherwerb unterstützen die Annahme, dass viel verständlicher Fremdspracheninput durch die Lehrkraft (Input-Hypothese, Krashen 1985; Bestandteil des Monitormodells) sowie häufige Interaktionen mit der Fremdsprache (*interaction approach*, Gass und Mackey 2007) auf einem ausreichend hohen Sprachniveau (Schwellenhypothese, Cummins 1979) die sprachlichen Fähigkeiten der Schülerinnen und Schüler besonders gut fördert. Diese Bedingungen scheinen im bilingualen Unterricht gegeben, da die Lerner für gewöhnlich einer vergleichsweise großen Menge an Fremdspracheninput durch die Lehrkraft ausgesetzt sind, der sich typischerweise auf einem etwas höheren als ihrem eigenen Sprachniveau befindet (vgl. Krashen 1985). Weiterhin bestehen ausreichend Möglichkeiten für die Lerner, mit der Fremdsprache zu interagieren und sie zu produzieren, um ihre Hypothesen zum korrekten Sprachgebrauch testen und mithilfe von Feedback anpassen zu können (vgl. Gass und Mackey 2007). Folglich könnten sich die Fremdsprachenkompetenzen der bilingual Unterrichteten auf einem ausreichend hohen Niveau befinden, um zusätzliche positive Effekte auf die Entwicklung der fremdsprachlichen Fähigkeiten zu ermöglichen; Effekte, die sich bei einer bloßen Überschreitung der unteren Kompetenzschwelle (gute Muttersprachen- und lediglich ausreichende Fremdsprachenfähigkeiten) nicht zeigen dürften (vgl. Cummins 1979). Weiterhin scheint die Überschreitung dieser oberen Schwelle (gute Muttersprachen- und Fremdsprachenkompetenzen) auch zu besseren Leistungen der Lerner im bilingual unterrichteten Sachfach führen zu können, da diese Überschreitung oft in gesteigertem „metalinguistic, academic, and cognitive functioning“ (Cummins 2006, S. 182) resultiert. Diesen kognitiven Zugewinn führen Adesope et al. (2010) auf den kontinuierlichen Wechsel der Lerner zwischen den Sprachen zurück, weshalb sich diese in der Regel sprachlich sowie kognitiv aufmerksamer und flexibler zeigten und Informationen nachhaltiger verarbeiteten als Lerner eines einsprachigen Settings.

Bisherige Forschung zur Sprachverwendung im bilingualen Unterricht

Die Anzahl der Untersuchungen zum Einsatz der Sprachen im bilingualen Unterricht ist sehr überschaubar. In einigen wenigen Studien wurden Lehrkräfte per Lehrerfragebogen oder in Interviews gebeten, hierzu Angaben zu machen. In DEZIBEL (Zydati 2007, S. 96) gaben diese an, dass die Verwendung des Deutschen nicht vollständig zugunsten der Fremdsprache unterdrckt werden sollte. In Mller-Schnecks (2006, S. 174) Studie erklrten lediglich 4% der 89 befragten

Lehrkräfte, vollständig auf den Einsatz von Deutsch zu verzichten. Ähnlich berichtete Kollenrott (2008, S. 138) von einer hohen, jedoch nicht 100%igen Zustimmung der 134 Befragten zur Aussage „Mein bilingualer Unterricht verläuft auf Englisch“. Bei Bredenbröker (2000, S. 103) gaben 89% der 25 Lehrkräfte an, Deutsch zumindest in Ausnahmefällen bzw. selten einzusetzen, die restlichen 11% sogar häufiger. Darüber hinaus erklärten 51% der 78 befragten Lehrkräfte in der Studie von Gierlinger et al. (2007, S. 78), dass auch die Schülerinnen und Schüler mindestens 90% aller Kommunikation auf Englisch abhielten. Weitere Studien gaben Auskunft darüber, in welchen Situationen Deutsch im bilingualen Unterricht eingesetzt wurde. Demnach verwandten die befragten Lehrkräfte die Muttersprache vor allem zweckgebunden, also in bestimmten Situationen zur Erreichung bestimmter Ziele. Es wurden folgende konkrete Situationen genannt: zur Erklärung komplizierter Sachverhalte, zur Absicherung des Wissens durch Wiederholungen und Zusammenfassungen, bei Einführung neuer Fachbegriffe (Bredenbröker 2000; Gierlinger et al. 2007), bei inhaltlichen und sprachlichen Schwierigkeiten (Müller-Schneck 2006) sowie bei Verständnisschwierigkeiten und Zeitmangel (Viebrock 2007).

Zusammenfassend liegen nur vereinzelte, wenig differenzierte und lediglich deskriptive empirische Befunde zur Rolle der Sprachen im bilingualen Unterricht vor. Bezüglich der Einführung von Fachtermini fehlen noch Informationen dazu, in welchem Umfang oder welcher Regel folgend diese zwei- oder einsprachig unterrichtet werden. Dies überrascht, da Fachbegriffe eine zentrale Rolle im bilingualen Unterricht einnehmen (Heimes 2013, S. 349) und in beiden Sprachen eingeführt werden sollten (KM 2006, S. 9; Otten und Wildhage 2007, S. 19; Theis 2010, S. 52). Schließlich sind keine Studien bekannt, die die Sprachverwendung in Bezug gesetzt haben zu den Leistungen bilingual unterrichteter Lerner (Hollm 2011, S. 329). Somit bleibt die Wirksamkeit solcher Unterrichtsaspekte bis dato ein Desiderat.

Forschungsfragen

Angesichts dieser Forschungslage bedarf es Studien, die Art und Umfang des Einsatzes von Mutter- und Fremdsprache analysieren sowie dessen Effekte auf die Entwicklung von Sprach- und Sachfachleistungen der Lerner untersuchen. Dies scheint speziell für das sprachbasierte Fach Geschichte lohnenswert. Weiterhin sollten mögliche Unterschiede zwischen den Klassen, beispielsweise im Vorkenntnisniveau oder in der Motivation, konstant gehalten werden, um unterschiedlich verlaufende Leistungszuwächse aufgrund solcher Faktoren ausschließen zu können. Dies führt zu folgenden zwei Forschungsfragen: Wie gestaltet sich die Verwendung von Mutter- und Fremdsprache im bilingualen Geschichtsunterricht? Die Beantwortung dieser Frage erfolgt im Hinblick auf bisherige Forschungsergebnisse anhand folgender Aspekte: Prozentuale Häufigkeit des Einsatzes von Englisch versus Deutsch durch Lehrkraft und Lerner, zweckgebundene Wahl der Sprache (Klärung komplexer Sachverhalte, Wiederholungen und Zusammenfassungen, Fachbegriffe) sowie Sprachwahl für die Einführung von Fachbegriffen. Es wird erwartet, dass Lehrkräfte und Lerner den Großteil bilingualen Unterrichts auf Englisch bestreiten und nur selten

die Muttersprache verwenden. Folglich wird weiter angenommen, dass Deutsch zwar in Abhängigkeit verschiedener Unterrichtssituationen eingesetzt wird, jedoch nur in sehr begrenztem Ausmaß. Bezüglich der Fachbegriffe wird erwartet, dass diese überwiegend in beiden Sprachen eingeführt werden.

Die zweite Forschungsfrage lautet: Welche Effekte zeigen diese sprachbezogenen Unterrichtsmerkmale auf den Leistungszuwachs der Lerner in Englisch und Geschichte unter Kontrolle von mittlerem Vorkenntnisniveau, sozioökonomischem Status und Motivation der Klassen? In Bezug auf theoretische Ansätze des Zweitspracherwerbs (vgl. Cummins 1979; Gass und Mackey 2007; Krashen 1985) wird erwartet, dass Lerner, deren Lehrkraft und die selbst im Unterricht auf einem ausreichend hohen Niveau häufig Englisch sprechen, einen größeren Zuwachs in den Englisch- und Geschichtsleistungen zeigen als Lerner, die selbst und deren Lehrkraft weniger Englisch verwenden. Weiterhin nehmen wir in Anlehnung an Butzkamms Prinzip der aufgeklärten Einsprachigkeit (1973) an, dass die Leistungszuwächse in Englisch und Geschichte dann besonders groß sind, wenn die Muttersprache häufig zweckgebunden eingesetzt wird. Aus diesem Grund erwarten wir eine ungünstigere Leistungsentwicklung der Lerner, wenn Fachbegriffe lediglich auf Englisch eingeführt werden. Wird hingegen die Wahl der Sprache reflektiert und in Abhängigkeit vom jeweiligen Begriff getroffen, oder werden diese grundsätzlich immer in beiden Sprachen unterrichtet, erwarten wir größere Leistungsentwicklungen.

Methode

Stichprobe und Studiendesign

Die Schülerstichprobe von COMBIH umfasste 1806 bilingual und einsprachig unterrichtete Achtklässler aus Baden-Württemberg. Für die vorliegende Arbeit sind jedoch nur die Daten der 703 bilingual Unterrichteten aus 30 Klassen (28 Gymnasien) relevant (50,1% weiblich; $M_{\text{Alter}} = 13,4$; $SD_{\text{Alter}} = 0,48$). 30,5% dieser Lerner besaßen Migrationshintergrund, d. h. mindestens ein Elternteil wurde nicht in Deutschland geboren. Bei 74% der Lerner verfügte mindestens ein Elternteil über die fachgebundene oder allgemeine Hochschulreife. Die Lerner waren Teil eines bilingualen Zugs in ihren Schulen. In Klassen 5 und 6 wurden sie zwei Schulstunden pro Woche zusätzlich in Englisch unterrichtet. Ab Klasse 7 nahmen sie in einem Fach pro Schuljahr an deutsch-englischem Sachfachunterricht teil. In Klasse 7 war dies das Fach Geographie, in der 8. Klasse wurde Geschichte zum ersten Mal bilingual unterrichtet. An deutschsprachigem Geschichtsunterricht nahmen die Lerner bereits in den Klassen 6 und 7 teil. Die Lehrerstichprobe umfasste 29 Geschichtslehrkräfte, die die bilingualen Klassen unterrichteten (41,4% weiblich; $M_{\text{Alter}} = 40,3$; $SD_{\text{Alter}} = 9,12$). Eine Lehrkraft nahm nicht an der Befragung teil.

Die Daten wurden im Herbst 2012 und Sommer 2013, zu Beginn und am Ende der 8. Klasse, erhoben. Die Datenerhebungen fanden in Stillarbeit während der regulären Unterrichtszeit statt und

wurden von Testleiterinnen mithilfe eines standardisierten Manuals durchgeführt. Die Lehrerfragebögen wurden entweder in der Schule oder zuhause ausgefüllt.

Erhebungsinstrumente

Sprachverwendung. Der prozentuale Anteil, zu dem Lehrkraft und Lerner die englische Sprache zur Kommunikation im Unterricht verwenden, wurde durch folgendes Item erhoben: „Bitte schätzen Sie: Wie viel Prozent sprechen Sie/ Ihre Schüler auf Englisch in einer durchschnittlichen bilingualen Geschichtsstunde?“ Die Lehrkraft sollte je eine Prozentzahl angeben. Dieses Item sowie alle folgenden Items zur Sprachverwendung wurden selbst verfasst.

Der zweckgebundene Einsatz des Deutschen wurde anhand des folgenden Items erfragt: „Wie oft sprechen Sie Deutsch in den folgenden Situationen im bilingualen Geschichtsunterricht dieser Klasse?“ Drei Optionen waren vorgegeben: „Zur (Er-)Klärung inhaltlich komplizierter Sachverhalte“, „Zur Wiederholung und Zusammenfassung von Unterrichtsstoff“, „Um Fachbegriffe auch auf Deutsch einzuführen“. Die Beantwortung erfolgte anhand einer vierstufigen Skala von 1 (*sehr selten – 1x pro Monat oder weniger*) bis 4 (*sehr häufig – mehr als 2x pro Schulstunde*).

Das Item zur Sprachwahl bei der Einführung von Fachbegriffen lautete: „Welche Aussage beschreibt am besten Ihren Umgang mit Fachbegriffen im bilingualen Geschichtsunterricht dieser Klasse?“ Die Lehrkräfte sollten sich für eine von fünf Antworten entscheiden: „Ich führe Fachbegriffe meistens nur auf Englisch ein“, „Ich führe Fachbegriffe meistens nur auf Deutsch ein“, „Ich führe Fachbegriffe meistens sowohl auf Englisch als auch auf Deutsch ein“, „Ich habe hierfür keine feste Regel, das hängt vom Begriff ab“, „Ich habe hierfür keine feste Regel, ich entscheide das meist spontan in der Stunde“. Die Items wurden später in Dummy-Variablen umkodiert (Referenzkategorie: *keine feste Regel, hängt vom Begriff ab*).

Englischleistungen. Die allgemeine englische Sprachfähigkeit wurde mithilfe eines selbst entwickelten C-Tests (vgl. Grotjahn 2011) erhoben. Er bestand aus 159 Items (d. h. 159 fehlende Worthälften in jedem zweiten Wort), die sich auf sieben Textpassagen mit unterschiedlichen Schwierigkeitsgraden verteilten. Um Testwiederholungseffekte zu reduzieren, wurde ein Booklet-Design verwandt: Zum ersten Messzeitpunkt (t1) füllten alle Lerner Texte 1, 3 und 5 aus. Die Hälfte der Lerner erhielt zusätzlich Texte 2 und 4, die andere Hälfte Texte 6 und 7. Zum zweiten Messzeitpunkt (t2) füllten alle Lerner wieder Texte 1, 3 und 5 sowie die beiden noch unbearbeiteten Texte aus. Zur psychometrischen Analyse der Items und der IRT-basierten Skalierung des Tests wurde ein 2-parametrisches Messmodell (Birnbaum-Modell) gewählt, das neben der Variation der Itemschwierigkeitsparameter auch unterschiedliche Trennschärfekoeffizienten der Items erlaubte (Birnbaum 1968). Die Items wurden im Hinblick auf *Differential Item Functioning* (DIF; Variation der Itemparameter zwischen verschiedenen Subpopulationen) sowie *Item Parameter Drift* (IPD; Variation der Itemparameter über die Zeit) untersucht, um unverfälschte Leistungsschätzungen zu gewährleisten. Für die DIF-Analyse wurden Geschlecht und die Gruppenzugehörigkeit (bilingual vs. einsprachig) berücksichtigt. Die Differenzen zwischen gruppen- und zeitpunktspezifischen

Itemparametern wurden unter Verwendung des Delta-Theorems auf statistische Signifikanz geprüft. Zusätzlich erfolgte für jedes Item eine Abschätzung der praktischen Signifikanz, indem die modellimplizierten itemcharakteristischen Kurven (*Item Characteristic Curves*; ICCs) der Subgruppen verglichen wurden. Unterschieden sich diese für die Gruppen deutlich, so war dies ein Hinweis auf DIF und das Item wurde entfernt. Dies traf auf 21 Items des Pretests sowie 18 Items des Posttests zu. Bezüglich der Drift-Analyse zeigten die teststatistische Prüfung der Itemparameter sowie die graphische Inspektion der ICCs, dass es deutliche messzeitpunktbezogene Unterschiede bei zwei Items gab. Folglich wurde für diese Items eine Variation der Itemparameter zwischen t1 und t2 (*Partial Measurement Invariance*) zugelassen. Danach wurde das finale Messmodell spezifiziert, das zur Schätzung der Personenparameter diente. Als Personenparameterschätzer wurden EAP-Schätzer verwendet. Zur Bestimmung der Reliabilität $Rel(EAP)$ der Tests wurde die Varianz der Personenparameter zur Varianz der latenten Fähigkeitsdimension Theta ins Verhältnis gesetzt. Die Reliabilität war zu beiden Messzeitpunkten hoch mit $Rel(EAP)t1 = 0,94$ und $Rel(EAP)t2 = 0,95$.

Weiterhin wurde das englische Hörverstehen der Lerner (Items vgl. Köller, Knigge und Tesch 2010) erhoben. Der Test bestand aus neun Aufgaben mit 43 offenen und geschlossenen Fragen. Zu t1 wurden Aufgaben 1 bis 6 eingesetzt. Zu t2 wurden Aufgaben 4 und 5 erneut verwandt (sog. Ankeritems) sowie um drei neue, etwas anspruchsvollere Aufgaben ergänzt, um dem Lernfortschritt der Schülerinnen und Schüler Rechnung zu tragen. Es wurde erneut ein 2-parametrisches Birnbaum-Modell angewandt (Birnbaum 1968). Geschlecht und Gruppenzugehörigkeit wurden als DIF-Variablen berücksichtigt und es wurden Drift-Analysen durchgeführt. Die ICCs von vier t1-Items unterschieden sich deutlich und wurden ausgeschlossen, während sich für die t2-Items kein DIF zeigte. Bezüglich der Drift-Analyse der Ankeritems zeigte sich, dass es keine nennenswerten Abweichungen hinsichtlich t1 und t2 gab. Deshalb wurden alle Parameter der Ankeritems zwischen Pre- und Posttest gleichgesetzt. Danach wurde das finale Messmodell spezifiziert. Als Personenparameterschätzer wurden EAP-Schätzer verwendet. Es ergab sich eine zufriedenstellende Reliabilität mit $Rel(EAP)t1 = 0,71$ und eine etwas höhere Reliabilität bei t2 mit $Rel(EAP)t2 = 0,84$.

Geschichtsleistungen. Das Geschichtswissen aus der 8. Klasse wurde mithilfe eines Lückentexts und Multiple-Choice-Fragen erhoben. Der Test wurde auf Grundlage des Baden-Württembergischen Bildungsplans entwickelt und umfasste 36 Items. Die Lerner mussten Auskunft zu wichtigen Daten, Namen und Fachbegriffen geben sowie korrekte historische Zusammenhänge, Gründe für und Konsequenzen von historischen Ereignissen identifizieren. Derselbe Test wurde zu Beginn und Ende des Schuljahres eingesetzt, um den Lernfortschritt messen zu können. Zu t1 erhielten ihn alle Lerner auf Deutsch. Zu t2 erhielt ihn ein Drittel der bilingual Unterrichteten auf Deutsch, die restlichen bilingual Unterrichteten erhielten entweder den Lückentext oder die Multiple-Choice-Fragen auf Englisch. Es wurde wieder ein 2-parametrisches Birnbaum-Modell gewählt

(Birnbaum 1968). Um mögliche Verzerrungen der Parameterschätzung durch die Variation der Sprache zu vermeiden, wurden bei der Skalierung die englischsprachigen Items als separate Items im Messmodell behandelt, deren Itemparameter sich von denen der entsprechenden deutschen Items unterscheiden durften. Es wurden DIF und Drift-Analysen durchgeführt. DIF zeigte sich bei drei t1-Items sowie bei zwei deutschsprachigen und zwei englischsprachigen t2-Items. Diese Items wurden ausgeschlossen. Bezüglich der Drift-Analyse zeigte sich, dass sich die Parameter von acht Items deutlich zwischen t1 und t2 unterschieden. Folglich wurde für diese Items eine Variation der Itemparameter zwischen t1 und t2 (*Partial Measurement Invariance*) zugelassen. Danach wurde das finale Messmodell spezifiziert. Als Personenparameterschätzer wurden EAP-Schätzer verwendet. Die Reliabilität war mit $Rel(EAP)t1 = 0,70$ zufriedenstellend und verbesserte sich bei t2 auf $Rel(EAP)t2 = 0,79$.

Kognitive Grundfähigkeiten. Zu t1 bearbeiteten die Lerner den figuralen und verbalen Subtest des Kognitiven Fähigkeitstests (KFT; Heller und Perleth 2000). Es wurden parallele Testversionen eingesetzt, um Abschreiben zu verhindern. Der figurale Test umfasste 25 Items mit Reliabilitäten von Cronbachs $\alpha = 0,91$ bzw. $\alpha = 0,90$ für Versionen A und B. Der verbale Test umfasste 20 Items mit $\alpha = 0,69$ bzw. $\alpha = 0,72$.

Motivation. Die intrinsische und extrinsische Motivation der Lerner in Englisch und Geschichte wurde, basierend auf dem Erwartungs-Mal-Wert-Modell von Eccles (vgl. Wigfield und Eccles 2000), mithilfe zweier Skalen erhoben: dem intrinsischen Wert (fünf Items; Geschichte: $\alpha = 0,82$; Englisch: $\alpha = 0,72$; z. B.: „Englisch/ Geschichte ist eines meiner Lieblingsfächer“) und dem Nützlichkeitswert (fünf Items; Geschichte: $\alpha = 0,89$; Englisch: $\alpha = 0,88$; z. B.: „Ich meine, dass man den Stoff in Englisch/ Geschichte auch später gut gebrauchen kann“). Antworten wurden auf einer vierstufigen Likert-Skala von 1 (*stimmt gar nicht*) bis 4 (*stimmt genau*) gegeben.

Familiärer Hintergrund. Die Items zur Erfassung des familiären Hintergrunds der Lerner wurden von large-scale Studien wie PISA (vgl. Klieme et al. 2010) adaptiert. Mit Schülerangaben zum Beruf und Schulabschluss der Eltern wurde für jeden Lerner das internationale sozioökonomische Maß des beruflichen Status (HISEI; Ganzeboom et al. 1992) bestimmt, der sich zwischen 11 und 88 (sehr niedriger bis sehr hoher sozioökonomischer Status) bewegte. Außerdem wurden Informationen zum Geschlecht und Migrationshintergrund (mindestens ein Elternteil nicht in Deutschland geboren) erfragt.

Statistisches Vorgehen

Die Variablen zur Sprachverwendung, die in COMBIH mittels Lehrerfragebögen erhoben wurden, beziehen sich immer auf die ganze Klasse, weshalb diese in Zweiebenenregressionsmodellen als Klassenmerkmale berücksichtigt wurden (Raudenbush und Bryk 2002). Es wurden für jeden Test (C-Test, Hörverstehen, Geschichtstest) vier Modelle in Mplus 7.1 (Muthén und Muthén 1998-2012) berechnet. Abhängige Variable war dabei jeweils die Testleistung der Lerner zu t2. Als unabhängige Variablen wurden auf Schülerebene (L1) Geschlecht, Migrationshintergrund, sozioökonomischer

Status, kognitive Grundfähigkeiten sowie Motivation und Vorwissen in Englisch und Geschichte berücksichtigt, um zu zeigen, welche Schülermerkmale mit einer besonders positiven Entwicklung im bilingualen Unterricht einhergingen. Die Merkmale der Sprachverwendung wurden als Prädiktoren auf Klassenebene (L2) berücksichtigt. In Modell 1 wurden die Variablen zum prozentualen Englischesinsatz in einer Schulstunde untersucht, in Modell 2 die zweckgebundene Verwendung des Deutschen und in Modell 3 die Variablen zur Sprachwahl bei Einführung von Fachbegriffen. Die Analysen wurden auf drei separate Modelle aufgeteilt, da sich die Daten auf L2 lediglich auf eine kleine Stichprobe ($N = 30$) bezogen. In allen Modellen wurde die durchschnittliche Vortestleistung der Klasse kontrolliert, um die Vorhersagekraft der Sprachverwendung für die Leistungsentwicklung zu bestimmen und eine mögliche Konfundierung von Vorleistungsniveau und Spracheinsatz zu kontrollieren. Weiterhin wurden der durchschnittliche sozioökonomische Status und die klassenbezogene intrinsische Motivation berücksichtigt, da diese Variablen ebenfalls substantielle Variationen zwischen den bilingualen Klassen aufwiesen (Intraklassenkorrelationskoeffizient $ICC \geq 0,05$). In Modell 0 wurden nur die Kontrollvariablen berücksichtigt, um zu bestimmen, wie viel zusätzliche Varianz durch die sprachbezogenen Variablen aufgeklärt wurde. Vor den Analysen wurden alle kontinuierlichen Variablen z-standardisiert und alle kontinuierlichen unabhängigen Variablen am Gruppenmittelwert zentriert.

Der Anteil fehlender Werte bewegte sich zwischen 0% (Geschlecht) und 13% (sozioökonomischer Status). Zur Schätzung dieser fehlenden Werte wurde das in Mplus 7 integrierte *full information maximum likelihood*-Verfahren (FIML) eingesetzt, das beim Schätzvorgang alle zur Verfügung stehenden Parameter berücksichtigt.

Ergebnisse

Zur Beantwortung der ersten Forschungsfrage zu Art und Umfang des Einsatzes von Mutter- und Fremdsprache im bilingualen Geschichtsunterricht berichtet Tab. 1 die Mittelwerte und Standardabweichungen bzw. prozentualen Häufigkeiten der Variablen zur Sprachverwendung. Es zeigte sich erwartungsgemäß, dass Lehrkräfte (90,3%) und Lerner (77,3%) überwiegend Englisch anstatt Deutsch zur Kommunikation verwendeten. Ebenso erwartungsgemäß verwiesen die niedrigen Mittelwerte der Variablen zum zweckgebundenen Einsatz des Deutschen auf einen eher seltenen Einsatz der Muttersprache. Wurde diese verwendet, dann am häufigsten zur Klärung inhaltlich komplexer Sachverhalte ($M = 2,03$; $SD = 0,65$). Wie erwartet gab ein Großteil der Lehrkräfte (41%) an, neue Fachtermini in beiden Sprachen einzuführen. 36,1% der Befragten unterrichteten Fachbegriffe nur in Englisch, während knapp ein Viertel angab, die Sprachwahl in Abhängigkeit vom jeweiligen Begriff zu treffen. Keine Lehrkraft gab an, Fachtermini lediglich auf Deutsch einzuführen oder spontan in der Schulstunde darüber zu entscheiden, ohne einer Regel zu folgen. Folglich wurden diese Antwortoptionen nicht als Variablen in die Regressionsanalysen aufgenommen.

Tab. 1: Mittelwerte und Standardabweichungen bzw. prozentuale Häufigkeiten der Variablen zur Sprachverwendung

| Variablen | <i>M</i> | <i>SD</i> |
|---|----------|-----------|
| Prozentualer Anteil Engischeinsatz in einer Unterrichtsstunde | | |
| Lehrkraft | 90,33 | 7,31 |
| Lerner | 77,29 | 13,60 |
| Häufigkeit des zweckgebundenen Einsatzes von Deutsch | | |
| (Er-)Klärung komplexer Inhalte | 2,03 | 0,65 |
| Wiederholung und Zusammenfassungen | 1,34 | 0,59 |
| Fachbegriffe | 1,95 | 0,91 |
| Sprachwahl beim Einführen von Fachbegriffen | | |
| Englisch | 36,1% | - |
| Deutsch | 0% | - |
| Englisch und Deutsch | 41,0% | - |
| Keine Regel, abhängig vom Begriff | 22,9% | - |
| Keine Regel, spontane Entscheidung | 0% | - |

Anmerkung. *M* = Mittelwert, *SD* = Standardabweichung.

Vor Beantwortung der zweiten Forschungsfrage zu den Effekten der Sprachverwendung auf die Leistungsentwicklung der Lerner im bilingualen Geschichtsunterricht werden zunächst mögliche Zusammenhänge zwischen individuellen Schülermerkmalen und der Leistungsentwicklung auf L1 präsentiert. Tab. 2 berichtet Regressionskoeffizienten, Standardfehler, Signifikanzen sowie die aufgeklärten Varianzanteile. Anschließend werden in Tab. 3 (C-Test, Hörverstehen) und Tab. 4 (Geschichtstest) die Ergebnisse der jeweils vier Regressionsmodelle zu den Effekten sprachbezogener Unterrichtsaspekte auf L2 vorgestellt.

Tab. 2 verweist auf eine Stabilität der Leistung im C-Test in Höhe von $B = 0,58$. Die Leistungsentwicklung im C-Test wurde von Geschlecht, figuralen und verbalen Fähigkeiten sowie dem Nützlichkeitswert vorhergesagt. D. h., bei Konstanthaltung aller jeweilig anderen Prädiktoren verlief die Leistungsentwicklung bei Mädchen positiver als bei Jungen sowie positiver bei Lernern mit hohen kognitiven Grundfähigkeiten und bei jenen, die Englisch als besonders nützlich empfanden. Bezüglich des Hörverstehenstests lag die Stabilität $B = 0,28$. Hier gingen die Variablen sozioökonomischer Status, verbale kognitive Fähigkeiten und Nützlichkeitswert mit dem Leistungszuwachs der Lerner einher. Beim Geschichtstest zeigte sich eine Stabilität von $B = 0,39$. Die Leistungsentwicklung wurde durch verbale Fähigkeiten und den intrinsischen Wert für Geschichte vorhergesagt.

Tab. 2: Effekte individueller Schülermerkmale auf die Leistungsentwicklung (Schülerebene)

| Variablen | C-Test | | Hörverstehen | | Geschichtstest | |
|--------------------------------|----------|---------------|--------------|---------------|----------------|---------------|
| | <i>B</i> | (<i>SE</i>) | <i>B</i> | (<i>SE</i>) | <i>B</i> | (<i>SE</i>) |
| Testleistung t1 | 0,58 | ** (0,18) | 0,28 | *** (0,04) | 0,39 | *** (0,05) |
| Geschlecht (1 = weiblich) | 0,15 | * (0,08) | -0,04 | (0,08) | -0,13 | (0,07) |
| Migrationshintergrund (1 = ja) | 0,03 | (0,06) | -0,03 | (0,06) | -0,03 | (0,05) |
| HISEI | 0,06 | (0,04) | 0,09 | * (0,04) | 0,06 | (0,04) |
| KFT figural | 0,11 | ** (0,04) | 0,07 | (0,04) | 0,03 | (0,04) |
| KFT verbal | 0,15 | ** (0,05) | 0,19 | *** (0,05) | 0,11 | * (0,04) |
| Intrinsischer Wert | 0,06 | (0,04) | 0,07 | (0,05) | 0,08 | * (0,03) |
| Nützlichkeitswert | 0,07 | * (0,03) | 0,09 | * (0,04) | 0,05 | (0,04) |
| Residualvarianz | 0,51 | | 0,71 | | 0,58 | |
| R ² | 0,43 | | 0,20 | | 0,28 | |

Anmerkung. *B* = Regressionskoeffizient, *SE* = Standardfehler; t1 = erster Messzeitpunkt, HISEI = internationales sozioökonomisches Maß des beruflichen Status, KFT = Kognitiver Fähigkeitstest.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Bezüglich der sprachbezogenen Variablen auf L2 zeigt Tab. 3, dass die Leistungsanstiege im C-Test in denjenigen Klassen größer waren, deren Lehrkräfte im Unterricht viel Englisch sprachen ($B = 0,21$; Modell 1). Für die anderen sprachbezogenen Variablen zeigten sich keine statistisch signifikanten Zusammenhänge. Aufgrund der kleinen Stichprobengröße auf L2 ($N_{\text{Klassen}} = 30$) sollte jedoch die durch die Sprachvariablen zusätzlich aufgeklärte Varianz in die Interpretation der Effekte einbezogen werden. So klärte deren Berücksichtigung 28% bzw. 5% und 11% zusätzliche Varianz der Gesamtvarianz auf L2 auf (vgl. Modelle 1 bis 3 mit Modell 0). Die Gesamtvarianz betrug beim C-Test 10%. Weiterhin schnitten diejenigen Klassen im C-Test besonders gut ab, deren mittlere Motivation besonders hoch ausfiel ($B = 0,37$; Modell 1).

Auch beim Hörverstehen (vgl. Tab. 3) verlief die Leistungsentwicklung in denjenigen Klassen besonders positiv, deren Lehrkräfte viel Englisch sprachen ($B = 0,13$; Modell 1). Dahingegen ging ein hoher englischer Redeanteil der Lerner mit einer ungünstigeren Leistungsentwicklung der Klassen einher ($B = -0,15$). Dieser Befund deutete sich auch in einer weiteren Analyse an, in der nur der englische Redeanteil der Schülerinnen und Schüler berücksichtigt wurde. Weiterhin zeigten diejenigen Klassen eine positivere Leistungsentwicklung, in denen verstärkt Deutsch für das Einführen von Fachbegriffen eingesetzt wurde ($B = 0,15$; Modell 2). Ähnlich erfolgte in Klassen, deren Lehrkräfte angaben, Fachbegriffe nur auf Englisch einzuführen, ein kleinerer Leistungszuwachs ($B = -0,54$; Modell 3) als in denjenigen Klassen, deren Lehrkräfte angaben, keiner strikten Regel zu folgen, sondern die Sprache in Abhängigkeit vom jeweiligen Begriff zu wählen. Schließlich zeigte sich die regelmäßige Verwendung beider Sprachen bei der Einführung von Fachtermini als dem Leistungszuwachs im Hörverstehen hinderlich ($B = -0,37$). Die Berücksichtigung der sprachbezogenen Variablen konnte 18% bzw. 28% und 39% zusätzliche Varianz in Bezug auf die Gesamtvarianz auf L2 aufklären (vgl. Modelle 1 bis 3 mit Modell 0). Diese

betrug beim Hörverstehen 11%. In Modell 2 wurde die Entwicklung des Hörverstehens auch durch das Vorkenntnisniveau der Klasse vorhergesagt ($B = 0,49$).

Bezüglich des Geschichtstests (vgl. Tab. 4) zeigte sich erneut, dass sich die Leistungen in einer Klasse dann besonders günstig entwickelten, wenn deren Lehrkraft besonders viel Englisch im Unterricht sprach ($B = 0,25$; Modell 1), die Lerner selbst im Schnitt jedoch eher weniger die Fremdsprache nutzten ($B = -0,28$). Diese Befunde zeigten sich auch in weiteren Analysen, in denen der englische Redeanteil der Lehrkraft und derjenige der Schülerinnen und Schüler separat berücksichtigt wurde. Bezüglich der anderen Variablen zeigten sich keine statistisch signifikanten Effekte. Lediglich das Einführen von Fachbegriffen auf Deutsch und Englisch ging mit einer ungünstigeren Leistungsentwicklung der Klassen einher ($B = -0,43$), als wenn deren Lehrkräfte angaben, keiner bestimmten Regel zu folgen, sondern die Sprache in Abhängigkeit vom Begriff zu wählen. Die Berücksichtigung der sprachbezogenen Variablen klärte 4% bzw. 14% und 28% zusätzliche Varianz der Gesamtvarianz auf L2 auf (vgl. Modelle 1 bis 3 mit Modell 0). Die Gesamtvarianz betrug beim Geschichtstest 22%. Schließlich sagte in Modell 1 auch die durchschnittliche intrinsische Motivation für Geschichte die Leistungsentwicklung der Klasse vorher ($B = 0,49$).

Tab. 3: Effekte sprachbezogener Unterrichtsaspekte auf die Leistungsentwicklung in Englisch (Klassenebene)

| Variablen | C-Test | | | | | | | | Hörverstehen | | | | | | | |
|----------------------|----------|-------------|----------|-------------|----------|-------------|----------|-------------|--------------|-------------|----------|-------------|----------|-------------|----------|-------------|
| | Modell 0 | | Modell 1 | | Modell 2 | | Modell 3 | | Modell 0 | | Modell 1 | | Modell 2 | | Modell 3 | |
| | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> |
| Intercept | -0,09 | (0,08) | -0,07 | (0,07) | -0,09 | (0,08) | -0,01 | (0,13) | 0,03 | (0,07) | 0,04 | (0,07) | 0,05 | (0,07) | 0,37 | ***(0,10) |
| Testleistung t1 | 0,25 | (0,19) | 0,10 | (0,13) | 0,24 | (0,19) | 0,12 | (0,20) | 0,38 | (0,28) | 0,38 | (0,23) | 0,49 | *(0,19) | 0,27 | (0,24) |
| HISEI | 0,28 | (0,16) | 0,18 | (0,13) | 0,27 | (0,15) | 0,28 | (0,15) | 0,22 | (0,12) | 0,09 | (0,13) | 0,09 | (0,13) | 0,20 | (0,11) |
| Intrinsischer Wert | 0,14 | (0,21) | 0,37 | *(0,16) | 0,08 | (0,16) | 0,20 | (0,12) | 0,11 | (0,17) | 0,16 | (0,17) | -0,02 | (0,15) | 0,02 | (0,13) |
| ENGredeanteil Lehrer | | | 0,21 | ***(0,06) | | | | | | | 0,13 | *(0,06) | | | | |
| ENGredeanteil SuS | | | -0,09 | (0,05) | | | | | | | -0,15 | *(0,07) | | | | |
| DEU Erklärungen | | | | | 0,00 | (0,06) | | | | | | | 0,08 | (0,06) | | |
| DEU Wiederholungen | | | | | 0,00 | (0,06) | | | | | | | -0,09 | (0,05) | | |
| DEU Begriffe | | | | | 0,07 | (0,07) | | | | | | | 0,15 | *(0,07) | | |
| Begriffe ENG | | | | | | | -0,24 | (0,15) | | | | | | | -0,54 | ***(0,11) |
| Begriffe ENG & DEU | | | | | | | 0,02 | (0,16) | | | | | | | -0,37 | ** (0,12) |
| Residualvarianz | 0,06 | | 0,04 | | 0,06 | | 0,05 | | 0,07 | | 0,05 | | 0,04 | | 0,03 | |
| R ² | 0,39 | | 0,67 | | 0,44 | | 0,50 | | 0,31 | | 0,49 | | 0,59 | | 0,70 | |

Anmerkung. *B* = Regressionskoeffizient, *SE* = Standardfehler; t1 = erster Messzeitpunkt, HISEI = internationales sozioökonomisches Maß des beruflichen Status, ENG = Englisch, DEU = Deutsch, SuS = Schülerinnen und Schüler.

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

Tab. 4: Effekte sprachbezogener Unterrichtsaspekte auf die Leistungsentwicklung im Geschichtstest (Klassenebene)

| Variablen | Modell 0 | | Modell 1 ENGredeanteil | | Modell 2 DEU zweckgebunden | | Modell 3 Einführung Fachbegriffe | |
|----------------------|----------|-------------|---------------------------|-------------|----------------------------------|-------------|--|-------------|
| | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> | <i>B</i> | <i>(SE)</i> |
| Intercept | 0,09 | (0,09) | 0,11 | (0,09) | 0,09 | (0,09) | 0,41 | *(0,19) |
| Testleistung t1 | 0,45 | (0,25) | 0,30 | (0,23) | 0,43 | (0,26) | 0,43 | (0,23) |
| HISEI | 0,29 | (0,19) | 0,10 | (0,15) | 0,30 | (0,19) | 0,23 | (0,18) |
| Intrinsischer Wert | 0,21 | (0,30) | 0,49 | *(0,20) | 0,19 | (0,27) | 0,20 | (0,25) |
| ENGredeanteil Lehrer | | | 0,25 | ** (0,09) | | | | |
| ENGredeanteil SuS | | | -0,28 | ** (0,09) | | | | |
| DEU Erklärungen | | | | | -0,06 | (0,08) | | |
| DEU Wiederholungen | | | | | 0,00 | (0,06) | | |
| DEU Begriffe | | | | | 0,09 | (0,07) | | |
| Begriffe ENG | | | | | | | -0,40 | (0,24) |
| Begriffe ENG & DEU | | | | | | | -0,43 | *(0,19) |
| Residualvarianz | 0,15 | | 0,09 | | 0,14 | | 0,11 | |
| R ² | 0,31 | | 0,59 | | 0,35 | | 0,45 | |

Anmerkung. *B* = Regressionskoeffizient, *SE* = Standardfehler; t1 = erster Messzeitpunkt, HISEI = internationales sozioökonomisches Maß des beruflichen Status, ENG = Englisch, DEU = Deutsch, SuS = Schülerinnen und Schüler.

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

Diskussion

Es war das Ziel des vorliegenden Beitrags, die Rolle der Sprachen im bilingualen Geschichtsunterricht zu analysieren. Hierfür verwandte COMBIH ein längsschnittliches Design mit Datenerhebungen zu Beginn und Ende eines 8. Schuljahres. Die erste Forschungsfrage untersuchte Art und Umfang des Einsatzes von Mutter- und Fremdsprache im bilingualen Geschichtsunterricht. Die Ergebnisse bestätigten unsere Annahmen und konnten bisherige Befunde (vgl. Bredenbröcker 2000; Gierlinger et al. 2007; Kollenrott 2008; Müller-Schneck 2006; Zydatiß 2007) replizieren, wonach im bilingualen Unterricht zwar überwiegend die Fremdsprache verwendet wird, jedoch Deutsch sowohl seitens der Lehrkräfte als auch der Lerner Berücksichtigung findet. Diese Ergebnisse stehen in Übereinstimmung mit den Vorgaben der Bildungspläne der Länder, die den überwiegenden Einsatz der Fremdsprache im bilingualen Unterricht vorsehen (vgl. Deutscher Bildungsserver 1996-2014), und unterstützen die bisher übliche Schulpraxis, die dies zu großen Teilen so umgesetzt hat (Böing und Palmes 2012, S. 73; Diehr 2012, S. 18). Weiterhin konnte mithilfe der vorliegenden Analysen bestätigt werden, dass Lehrkräfte Deutsch zweckgebunden in verschiedenen Situationen einsetzen (vgl. Bredenbröcker 2000; Gierlinger et al. 2007; Viebrock 2007). Darüber hinaus konnten in COMBIH bisherige Befunde durch detaillierte Informationen zur Sprachwahl bei der Einführung

von Fachbegriffen ergänzt werden, wonach Fachtermini überwiegend nur auf Englisch oder in beiden Sprachen eingeführt wurden, jedoch nie nur auf Deutsch.

Darüber hinaus setzte COMBIH diese deskriptiven Befunde erstmals in Bezug zu Schülerleistungen. Die zweite Forschungsfrage hatte zum Ziel, die Wirksamkeit der sprachbezogenen Unterrichtsmerkmale hinsichtlich der Leistungsentwicklung der Klassen in Englisch und Geschichte zu analysieren. Die großen Varianzanteile, die durch die Berücksichtigung dieser Sprachvariablen zusätzlich aufgeklärt wurden, verweisen auf die hohe Relevanz dieses Zusammenhangs. Es wurden sowohl Effekte des Einsatzes der Fremd- als auch der Muttersprache gefunden. Bezüglich der Verwendung von Englisch konnten vorliegende Ergebnisse erwartungsgemäß zeigen, dass ein großer fremdsprachiger Redeanteil durch die Lehrkraft mit größeren Leistungsentwicklungen der Klassen im C-Test und im Hörverstehen einherging. Dasselbe traf auf die Entwicklung der Geschichtskennntnisse zu, die in denjenigen Klassen positiver verlief, deren Lehrkräfte viel Englisch sprachen. Somit scheinen Krashens Input-Hypothese (1985) sowie Cummins Schwellenhypothese (1979) auf den Kontext des bilingualen Unterrichts übertragbar zu sein. Bezüglich des Englischredeanteils der Lerner widersprachen die Befunde unseren Erwartungen. Ein starker Gebrauch des Englischen ging mit ungünstigeren Leistungsentwicklungen der Klassen einher. Dies könnte daran liegen, dass die mündlichen Schülerbeiträge möglicherweise zu viele fremdsprachenbezogene Fehler enthielten. Folglich hätten sich die produktiven Englischfähigkeiten der Lerner nicht auf einem ausreichend hohen Niveau befunden, um zusätzliche positive Effekte zu bewirken (vgl. Cummins 1979).

Dieser unerwartete Befund verweist auf die Wichtigkeit der Verwendung der Muttersprache im bilingualen Geschichtsunterricht. Denn in Anlehnung an Butzkamms Überlegungen (1973, 2002) wäre denkbar, dass der zu geringe Deutscheinsatz seitens der Lerner in nicht ausreichenden Fremdsprachenfähigkeiten sowie sachfachlichen Verständnisproblemen resultierte und somit günstigere Leistungsentwicklungen hemmte. Weiterhin scheint jedoch nicht nur der Einbezug der Muttersprache per se von Bedeutung zu sein, sondern vielmehr deren reflektierte und planvolle Verwendung. Darauf verweisen COMBIHs Ergebnisse, wonach bei Einführung von Fachbegriffen nicht nur die alleinige Verwendung von Englisch, sondern auch der durchgängige Einsatz beider Sprachen mit weniger positiven Entwicklungen des Hörverstehens und Geschichtswissens einherging. In Klassen hingegen, in denen die Lehrkraft angab, die Sprachwahl reflektiert in Abhängigkeit vom Begriff zu treffen, zeigten sich positivere Ergebnisse.

Zusammenfassend liegt die Bedeutung der vorliegenden Befunde darin, dass sie erstmals Aussagen zur Wirksamkeit von sprachbezogenen Unterrichtsmerkmalen hinsichtlich der Leistungsentwicklung von Schülerinnen und Schülern im bilingualen Geschichtsunterricht treffen konnten. Dies geschah mithilfe von Mehrebenenanalysen, welche die vorgelegene hierarchische Datenstruktur adäquat berücksichtigten. Jedoch konnte nur ein kleiner inhaltlicher Ausschnitt untersucht werden. Zukünftige Studien müssten daher die Rolle der Sprachen im bilingualen

Unterricht noch differenzierter untersuchen und hierbei deren Wirksamkeit hinsichtlich produktiver Fremdsprachenfähigkeiten sowie anderer Kompetenzen verschiedener Sachfächer in den Blick nehmen. Weiterhin sollte die Wirksamkeit weiterer methodisch-didaktischer Aspekte, beispielsweise Materialeinsatz oder Sozialformen, im bilingualen Unterricht analysiert werden. Die Unterrichtsbeobachtung wäre hierfür ein geeignetes Instrument, um unterrichtliche Geschehnisse noch präziser zu erfassen. Diese Ergebnisse könnten dann als Grundlage für die noch ausstehende Entwicklung einer spezifischen Methodik und Didaktik des bilingualen Unterrichts dienen, die zum Ziel hat, die Kompetenzen der Schülerinnen und Schüler möglichst optimal zu fördern. Die vorliegenden Ergebnisse deuten hierzu an, dass sowohl verstärkter englisch- als auch deutschsprachiger Input durch die Lehrkraft sowie muttersprachliche Äußerungen durch die Lerner von Bedeutung sein könnten.

Literaturverzeichnis

- Adesope, O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207–245.
- Badertscher, H., & Bieri, T. (2009). *Wissenserwerb im Content and Language Integrated Learning*. Bern: Haupt.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. Lord & M. Novick (Hrsg.), *Statistical theories of mental test scores* (S. 397–479). Reading, MA: Addison-Wesley.
- Böing, M., & Palmen, P. (2012). Bilingual heißt zweisprachig! Überlegungen zur Verwendung beider Sprachen im bilingual deutsch-französischen Geographieunterricht. In B. Diehr & L. Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken* (S. 73–90). Frankfurt/Main: Lang.
- Bonnet, A. (2007). Fach, Sprache, Interaktion – Eine Drei-Säulen-Methodik für CLIL. *Fremdsprachen Lehren und Lernen, 36*, 126–141.
- Bredenbröker, W. (2000). *Förderung der fremdsprachlichen Kompetenz durch bilingualen Unterricht*. Frankfurt/Main: Lang.
- Breidbach, S. (2010). Bilinguale Didaktik – bald wieder zwischen allen Stühlen? Zu den Aussichten einer integrativen Didaktik des bilingualen Sachfachunterrichts. In G. Bach & S. Niemeier (Hrsg.), *Bilingualer Unterricht* (S. 165–176). Frankfurt/Main: Lang.
- Butzkamm, W. (1973). *Aufgeklärte Einsprachigkeit. Zur Entdogmatisierung der Methode im Fremdsprachenunterricht*. Heidelberg: Quelle & Meyer.
- Butzkamm, W. (2002). *Psycholinguistik des Fremdsprachenunterrichts*. Tübingen: Francke.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222–251.
- Dallinger, S., Jonkmann, K., & Hollm, J. (2014, März). COMBIH: Kompetenzentwicklung im bilingualen Geschichtsunterricht. Posterpräsentation auf der GEBF in Frankfurt/Main.
- Deutscher Bildungsserver (1996–2014). *Bildungspläne der Bundesländer für allgemeinbildende Schulen*. <http://www.bildungsserver.de/Bildungsplaene-Lehrplaene-der-Bundeslaender-fuer-allgemeinbildende-Schulen-400.html>. Zugegriffen: 16. Dezember 2014.
- Diehr, B. (2012). What's in a name? Terminologische, typologische und programmatische Überlegungen zum Verhältnis der Sprachen im Bilingualen Unterricht. In B. Diehr & L. Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken* (S. 17–36). Frankfurt/Main: Lang.
- Eurydice (Hrsg.) (2006). *Content and Language Integrated Learning (CLIL) at school in Europe*. <http://bookshop.europa.eu/en/content-and-language-integrated-learning-clil-at-school-in-europe-pbNCX106001>. Zugegriffen: 19. September 2014.
- Ganzeboom, H., de Graaf, P., Treimann, D., & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*(1), 1–56.

- Gass, S., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Hrsg.), *Theories in second language acquisition* (S. 175–199). Mahwah, NJ: Erlbaum.
- Gierlinger, E., Hametner, C., & Spann, H. (2007). Englisch als Arbeitssprache – viele Inseln und (noch) kein Festland? In E. Gierlinger, V. Hainschink, & H. Spann (Hrsg.), *Content and Language Integrated Learning (CLIL) auf der Sekundarstufe I* (S. 63–118). Linz: Trauner.
- Grotjahn, R. (2011). C-Tests: Aspekte der Validität. *Deutsch als Fremdsprache*, 48(3), 131–137.
- Heimes, A. (2013). Didaktiken und Methodiken bilingualer Fächer: Geschichte. In W. Hallet & F. Königs (Hrsg.), *Handbuch Bilingualer Unterricht* (S. 345–352). Seelze: Kallmeyer.
- Heine, L. (2010). Fremdsprache und konzeptuelle Repräsentation. Bilingualer Unterricht aus kognitiver Perspektive. In S. Doff (Hrsg.), *Bilingualer Sachfachunterricht in der Sekundarstufe* (S. 199–212). Tübingen: Narr.
- Heller, K., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. - 12. Klassen. Revision*. Göttingen: Hogrefe.
- Hollm, J. (2011). Wortschatzarbeit im bilingualen Sachfachunterricht. In P. Schäfer & C. Schowalter (Hrsg.), *In mediam linguam. Mediensprache – Redewendungen – Sprachvermittlung* (S. 319–329). Landau: Verlag Empirische Pädagogik.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., & Prenzel, M. (Hrsg.) (2010). *PISA 2009*. Münster: Waxmann.
- KM (2006) = Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2006). *Bilingualer Unterricht. Realschule*. Braunschweig: Westermann.
- KM (2008) = Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2008). *Bilingualer Unterricht deutsch-englisch an allgemein bildenden Gymnasien*. Stuttgart: Schwäbische Druckerei.
- Kollenrott, A. (2008). *Sichtweisen auf deutsch-englisch bilingualen Geschichtsunterricht*. Frankfurt/Main: Lang
- Köller, O., Knigge, M., & Tesch, B. (Hrsg.) (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Köller, O., Leucht, M., & Pant, H. (2012). Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I. *Unterrichtswissenschaft*, 4(4), 334–350.
- Krashen, S. D. (1985). *The input hypothesis. Issues and implications*. London: Longman.
- Müller-Schneck, E. (2006). *Bilingualer Geschichtsunterricht*. Frankfurt/Main: Lang.
- Muthén, L., & Muthén, B. (1998–2012). *Mplus user's guide*. Los Angeles: Muthén & Muthén.
- Nold, G., Hartig, J., Hinz, S., & Rossa, H. (2008). Klassen mit bilingualem Sachfachunterricht. Englisch als Arbeitssprache. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch* (S. 451–457). Weinheim: Beltz.

- Otten, E., & Wildhage, M. (2007). Content and Language Integrated Learning. Eckpunkte einer „kleinen“ Didaktik des bilingualen Sachfachunterrichts. In M. Wildhage & E. Otten (Hrsg.), *Praxis des bilingualen Unterrichts* (S. 12–45). Berlin: Cornelsen.
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models*. Thousand Oaks: Sage.
- Rumlich, D. (2014). Prospective CLIL and non-CLIL students' interest in English (classes). In R. Breeze, C. Pasamar, C. Saíz, & C. Sala (Hrsg.), *Integration of theory and practice in CLIL* (S. 75–95). Amsterdam: Rodopi.
- Theis, R. (2010). Bilingualer Geschichtsunterricht. In S. Doff (Hrsg.), *Bilingualer Sachfachunterricht in der Sekundarstufe* (S. 44–57). Tübingen: Narr.
- Thürmann, E. (2013). Spezifische Methoden für den bilingualen Unterricht/CLIL. In W. Hallet & F. Königs (Hrsg.), *Handbuch Bilingualer Unterricht* (S. 229–235). Seelze: Kallmeyer.
- Viebrock, B. (2007). *Bilingualer Erdkundeunterricht. Subjektive didaktische Theorien von Lehrerinnen und Lehrern*. Frankfurt/Main: Lang.
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.
- Zydati, W. (2007). *Deutsch-Englische Zge in Berlin (DEZIBEL)*. Frankfurt/Main: Lang.

Abschlussdiskussion

7 Abschlussdiskussion

Das Ziel der vorliegenden Studie war es, die Wirksamkeit deutsch-englischen bilingualen Geschichtsunterrichts hinsichtlich der Schülerleistungen zu untersuchen. Dies scheint insbesondere deshalb relevant, da der bilinguale Sachfachunterricht im letzten Jahrzehnt deutschlandweit systematisch ausgebaut wurde und mittlerweile in allen Bundesländern in den verschiedensten Schulformen, Fächern und Fremdsprachen angeboten wird (KMK, 2013: 11-14). Informationen zur Wirksamkeit dieser Unterrichtsform geben Auskunft darüber, was bilingualer Unterricht bei seinen Schülerinnen und Schülern bewirkt und auf welche methodisch-didaktischen Vorgänge dies möglicherweise zurückgeführt werden kann. Entsprechende Befunde können somit als Grundlage für die zukünftige Organisation und Entwicklung dieser Unterrichtsform dienen. Zur Umsetzung dieses Forschungsvorhabens wurde als theoretische Basis das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts von Helmke (2012) gewählt, da es verschiedenste Anknüpfungspunkte für eine umfassende und möglichst adäquate Analyse der Wirksamkeit bilingualen Unterrichts bietet, wie beispielsweise „Familie“, „Lernpotential“ und „Klassenzusammensetzung“, „Lehrperson“, „Unterrichtsangebot“ und dessen „Wahrnehmung“ sowie „Ertrag“ (vgl. Abschnitte 1.2 bis 1.5). Hierfür erfolgte eine interdisziplinäre Zusammenarbeit zwischen Fremdsprachen- und Sachfachdidaktik sowie der empirischen Bildungsforschung, um den entsprechenden inhaltlichen Fragestellungen sowie forschungsmethodischen und statistisch-analytischen Anforderungen gerecht zu werden. Mithilfe dieser Vorgehensweise konnten in der vorliegenden Arbeit innovative Fragestellungen und Methoden für die Analyse der Wirksamkeit bilingualen Unterrichts aufgegriffen und eingesetzt werden, die zu belastbaren und neuen Befunden in diesem Bereich führten. Diese werden im Folgenden zusammengefasst und deren Bedeutung diskutiert. Hieran anknüpfend werden im darauffolgenden Abschnitt weitere wichtige Anschlussfragestellungen aufgezeigt, die die vorliegenden Befunde ergänzen und zu weiteren Erkenntnissen hinsichtlich der Wirksamkeit bilingualen Unterrichts verhelfen können.

7.1 Zusammenfassung und Bedeutung der zentralen Befunde

7.1.1 Zusammenfassung der zentralen Befunde

Die vorliegende Arbeit untersuchte die Wirksamkeit bilingualen Geschichtsunterrichts im Hinblick auf die Schülerleistungen in der Fremdsprache und im Sachfach. Hierfür wurden in einem ersten Schritt mögliche Unterschiede in den Lernausgangslagen zwischen bilingual und einsprachig unterrichteten Lernern analysiert, um diese in einem zweiten Schritt statistisch konstant halten und somit den Effekt dieser Unterrichtsform auf die Kompetenzentwicklung der Lerner bestmöglich isolieren zu können (vgl. Abschnitt 1.6.1). Diesbezüglich wurde in der vorliegenden Arbeit gezeigt,

dass zu Beginn des bilingualen Geschichtsunterrichts (was dem Beginn des vierten Schuljahrs im bilingualen Zug entspricht) substantielle Unterschiede zugunsten der bilingual unterrichteten Schülerinnen und Schüler bestanden. Dies traf sowohl auf deren Lernvoraussetzungen (sozioökonomischer Status, Bildungsabschluss der Eltern, kulturelles Kapital, verbale kognitive Grundfähigkeiten) als auch auf die Motivation und das Vorwissen in Geschichte zu. Dieser Befund kann vermutlich größtenteils auf die positive Selektion der Lerner zurückgeführt werden, die im Rahmen der Auswahlverfahren bilingualer Zug-Schulen stattzufinden scheint (vgl. Abschnitt 1.2.1). Kleinere Unterschiede in den Bereichen Schüलगeschlecht, Migrationshintergrund und figurale kognitive Grundfähigkeiten waren statistisch nicht signifikant. Weiterhin wurden in der vorliegenden Arbeit deutliche Vorteile in der Englischmotivation (intrinsisches Interesse, Wichtigkeit, Nützlichkeit, Kosten, akademisches Selbstkonzept) und in den Englischkompetenzen (allgemeine englische Sprachfähigkeit, Hörverstehen) zugunsten der bilingual Unterrichteten festgestellt, die mit großer Wahrscheinlichkeit durch eine personenimmanente sowie, durch die Teilnahme am bilingualen Zug herbeigeführte, gesteigerte Sprachaffinität zu begründen sind. Dieser Vorsprung in den Englischkompetenzen in Höhe von bis zu zwei Schuljahren verringerte sich zwar etwas durch die Berücksichtigung der günstigeren Lernausgangslagen der bilingual unterrichteten Schülerinnen und Schüler, er bestand aber dennoch in einem Umfang von mindestens eineinviertel Schuljahren weiter. Somit scheint der Selektionseffekt bilingualer Unterrichtsprogramme lediglich einen Teil der Unterschiede in den Englischleistungen zwischen bilingual und einsprachig unterrichteten Lernern aufzuklären. Ein weiterer Faktor, der jedoch in der vorliegenden Studie nicht berücksichtigt werden konnte, stellt der zusätzliche vorbereitende Englischunterricht in den Klassenstufen 5 und 6 als Teil des bilingualen Zugs dar (vgl. Abschnitt 1.2.1), der mit großer Wahrscheinlichkeit zu diesem Vorteil der bilingual unterrichteten Lerner in ihren Englischfähigkeiten beigetragen hat. Weitere Leistungsunterschiede, die nach Einbezug dieses Aspekts immer noch bestehen könnten, könnten dann auf einen positiven Effekt des bilingualen Unterrichts auf die Englischkompetenzen verweisen. Der Vergleich von Lernvoraussetzungen und Vorkenntnissen zwischen den beiden einsprachig unterrichteten Schülergruppen (Schülerinnen und Schüler aus den Parallelklassen der bilingual Unterrichteten sowie aus Schulen ohne bilinguales Unterrichtsangebot) zeigte keine statistisch signifikanten Unterschiede. Somit scheinen keine Nachteile für diejenigen Lerner zu bestehen, die sich an Schulen mit bilingualem Zug gegen diese Spezialisierung und für ein anderes Profil entschieden haben (Teilstudien I.A und I.B).

Weiterhin konnte gezeigt werden, dass die bilingual unterrichteten Klassen im Verlauf eines Schuljahres bilingualen Geschichtsunterrichts ihren Leistungsvorsprung im englischen Hörverstehen merklich weiter ausbauen konnten gegenüber den einsprachig unterrichteten Gruppen. Dies ist ein bemerkenswerter Befund angesichts der Tatsache, dass in den entsprechenden Analysen Unterschiede in den Lernausgangslagen und Vorkenntnisständen sowie in der Unterrichtsqualität zugunsten der bilingual unterrichteten Klassen statistisch kontrolliert wurden. Bezüglich der

Entwicklung der allgemeinen englischen Sprachfähigkeit zeigte sich diese größere Leistungsentwicklung jedoch nicht. Die zunächst stark positiven Effekte zugunsten der bilingual Unterrichteten am Ende des Schuljahres waren statistisch lediglich noch marginal bzw. nicht mehr statistisch signifikant, nachdem die verschiedenen Kontrollvariablen berücksichtigt wurden. Diese Ergebnisse verweisen in Zusammenschau auf eine tendenziell positive Wirkung des bilingualen Unterrichts auf die Englischkompetenzen der Lerner, insbesondere bezüglich des Hörverstehens, und können unter Zuhilfenahme verschiedener theoretischer Ansätze des Zweitspracherwerbs begründet werden (Gass & Mackey, 2007; Krashen & Terrell, 2000; Krashen, 1985). Diese deuten Vorteile beim Erwerb einer Fremdsprache an, wenn diesbezüglich genug verständlicher fremdsprachlicher Input erfolgt und die Lerner die Fremdsprache in einer natürlichen Spracherwerbsumgebung mit häufigen Gelegenheiten zur eigenen Sprachproduktion und -korrektur erlernen – Bedingungen, die im bilingualen Unterricht gegeben sind (vgl. Abschnitt 1.5.1). Bezüglich des Sachwissens in Geschichte zeigten sich vergleichbare Leistungsentwicklungen zwischen den Klassen – und zwar trotz der Tatsache, dass die bilingual Unterrichteten an 50% mehr Geschichtunterricht pro Woche teilnahmen als die beiden Vergleichsgruppen. Somit scheinen sich die kognitiven Vorteile, die bei bilingual aufwachsenden Kindern zu beobachten sind (Adesope, Lavin, Thompson & Ungerleider, 2010) bzw. beim Erwerb einer Fremdsprache in einer zweisprachigen Unterrichtssituation erwartet werden (Cummins, 1979), nicht auf den Kontext des „künstlichen“ bilingualen Unterrichtssettings der vorliegenden Arbeit übertragen zu lassen (vgl. Abschnitt 1.5.1). Dasselbe trifft auch auf den möglichen semantischen Mehrwert für das sachfachliche Wissen zu, bezüglich dessen angenommen wurde, dass dieser durch semantische Diskrepanzen in den Fachbegriffen der beiden Sprachen ausgelöst wird (vgl. Abschnitt 1.4.1). Zusammenfassend kann geschlussfolgert werden, dass sich bilingualer Unterricht positiv auf das englische Hörverstehen der Klassen auszuwirken scheint, während sich für die allgemeine englische Sprachfähigkeit und das Geschichtswissen kein nennenswerter Effekt dieser Unterrichtsform auf die Schülerleistungen zeigte (Teilstudie II).

Abschließend wurde in der vorliegenden Arbeit die Wirksamkeit sprachbezogener methodisch-didaktischer Aspekte des bilingualen Unterrichts hinsichtlich der Leistungen der Klassen in Englisch und Geschichte aufgezeigt. Hierbei wurden mögliche Unterschiede in den Vorkenntnisständen sowie im sozioökonomischen Status und der Motivation zwischen den Klassen statistisch kontrolliert. Die Ergebnisse verwiesen auf Zusammenhänge zwischen Art und Umfang der Verwendung von Mutter- und Fremdsprache und der Leistungsentwicklung der Klassen sowohl in Englisch als auch in Geschichte. Hierbei zeigten sich einerseits positive Effekte bei hohen englischen Redeanteilen seitens der Lehrkräfte. Dies kann in Verbindung mit Krashens (1985) Input-Hypothese und Cummins (1979) Schwellenhypothese gebracht werden, wonach ein ausreichend großer Fremdspracheninput auf einem etwas höheren als dem eigenen Niveau den Fremdspracherwerbsprozess der Lerner positiv beeinflussen und weiterhin zu Vorteilen im

Sachfach führen kann. Andererseits ging auch der reflektierte Miteinbezug der Muttersprache, insbesondere bei der Einführung neuer Fachtermini, mit günstigen Leistungsentwicklungen der Klassen im Hörverstehen und in Geschichte einher. Diesbezüglich kann auf die Annahmen von Butzkamms (1973) Prinzip der aufgeklärten Einsprachigkeit verwiesen werden, wonach der planvolle Einsatz der Muttersprache den Lernfortschritt in der Fremdsprache begünstigen sollte (vgl. Abschnitt 1.5.1). Schließlich zeigten die vorliegenden Analysen, dass die fast ausschließliche Verwendung von Englisch auf Kosten deutscher Redeanteile seitens der Schülerinnen und Schüler in schlechteren Leistungsentwicklungen zu resultieren schien. Dies stellt die Übertragung des *interaction approach* (Gass & Mackey, 2007) auf den Kontext des bilingualen Sachfachunterrichts in Frage. Dieser besagt, dass ausreichende und qualitativ angemessene Interaktionen mit der Fremdsprache sowie deren durch Feedback unterstützte Produktion seitens der Lerner zu einer Verbesserung der Fremdsprachenkompetenzen führen sollte (Teilstudie III).

7.1.2 Forschungsmethodisch bedingte Bedeutung der Befunde

Die Bedeutsamkeit der vorliegenden Befunde liegt darin, dass sie die bisherigen Ergebnisse zur Wirksamkeit bilingualen Unterrichts teilweise replizieren und diese darüber hinaus entscheidend differenzieren und ergänzen können (vgl. Abschnitt 1.6). Unter Bezugnahme auf Helmkes (2012) Angebots-Nutzungs-Modell wurde dies mithilfe der interdisziplinären Zusammenarbeit zwischen Englisch- und Geschichtsdidaktik sowie der empirischen Bildungsforschung und der Berücksichtigung der damit in Zusammenhang stehenden Fragestellungen und analytisch-forschungsmethodischen Vorgehensweisen möglich, die zum Ziel hatten, die Wirksamkeit bilingualen Unterrichts möglichst umfassend und adäquat zu untersuchen.

Mithilfe der Ergebnisse zu den Effekten dieser Unterrichtsform auf die Schülerleistungen in Englisch und Geschichte konnte beispielsweise die Wirksamkeit bilingualen Unterrichts sowohl hinsichtlich der Fremdsprachenfähigkeiten als auch des Geschichtswissens der Lerner festgestellt werden. Hierfür wurde auf die didaktische Expertise der Fächer Englisch und Geschichte zurückgegriffen und entsprechende Messinstrumente entwickelt bzw. ausgewählt. Bisherige Studien konnten einen solchen Doppelfokus kaum aufweisen, meist aufgrund fehlender Analysen zum Sachfach (Bonnet, 2012: 67; Bernhardt, 2011: 215; Coyle, Hood & Marsh, 2010: 138; Sylvén, 2010: 224). Weiterhin wurde für eine möglichst adäquate Untersuchung der Wirksamkeit bilingualen Unterrichts ein längsschnittliches Studiendesign gewählt. Dieses erfasste die Schülerleistungen von bilingual sowie einsprachig unterrichteten Schülerinnen und Schülern aus zwei Vergleichsgruppen zu zwei Messzeitpunkten: zu Beginn des Schuljahres (und somit auch zu Beginn des bilingualen Geschichtsunterrichts) und zum Ende dieses Schuljahres. Dieses Studiendesign, das die Vorkenntnisse der Lerner sowohl in der Fremdsprache als auch im Sachfach sowie den Vergleich mit zwei verschiedenartigen einsprachig unterrichteten Schülergruppen miteinbezog, wurde von

empirischen Bildungsforschern angeregt. Es stellt im Vergleich zu bisherigen Untersuchungen zur Wirksamkeit bilingualen Unterrichts eine Besonderheit dar.

Weiterhin wurden in der vorliegenden Arbeit verschiedenste Aspekte berücksichtigt, bezüglich derer Unterschiede zwischen bilingual und einsprachig unterrichteten Schülergruppen erwartet wurden und die mit großer Wahrscheinlichkeit die Leistungen der Klassen beeinflussten (Helmke, 2012), welche jedoch nicht in direktem Zusammenhang mit dem bilingualen Unterrichtsetting standen. Hierzu zählten Geschlecht, Migrationshintergrund, Bildungsnähe der Eltern (Schulabschluss, Beruf, kulturelles Kapital), sozioökonomischer Status, verbale und figurale kognitive Grundfähigkeiten, Motivation in Englisch und Geschichte (intrinsisches Interesse, Wichtigkeit, Nützlichkeit, Kosten, akademisches Selbstkonzept), Vorkenntnisse in Englisch und Geschichte, Unterrichtsqualität in Englisch und Geschichte (Klassenführung, konstruktive Unterstützung, kognitive Aktivierung) sowie Enthusiasmus und Selbstwirksamkeitsüberzeugung der Lehrkraft. Manche dieser Merkmale wurden in bisherigen Studien bereits einbezogen (z. B. verbale kognitive Fähigkeiten und sozioökonomischer Status; Rumlich, 2014b; Köller, Leucht & Pant, 2012; Nold, Hartig, Hinz & Rossa, 2008). Andere hingegen wurden in der vorliegenden Studie erstmals berücksichtigt (z. B. figurale kognitive Fähigkeiten, Vorkenntnisstand im Klassendurchschnitt, Unterrichtsqualität, Lehrervariablen). Die notwendigen Tests und Skalen zur Erfassung dieser Schüler- und Klassenmerkmale stammten überwiegend aus dem Bereich der empirischen Bildungsforschung bzw. Pädagogischen Psychologie.

Die eben genannten Aspekte wurden in der vorliegenden Studie deshalb so umfassend berücksichtigt, um die Effekte bilingualen Unterrichts auf die Leistungen der Klassen bestmöglich isolieren zu können. Somit sollte die Überschätzung dieser Effekte verhindert werden, was in anderen Studien zur Wirksamkeit bilingualen Unterrichts bisher oftmals der Fall gewesen war (Rumlich, 2014a: 83). Dies wurde weiterhin auch dadurch erreicht, dass in der vorliegenden Arbeit erstmals die Leistungsentwicklung ganzer bilingual unterrichteter Klassen in den Blick genommen wurde. Folglich konnte mithilfe von Mehrebenenregressionsanalysen (Raudenbush & Bryk, 2002) und der damit in Zusammenhang stehenden Berücksichtigung der geklumpten Stichprobe (Lerner in Klassen) der Tatsache Rechnung getragen werden, dass bilingualer Unterricht ein klassenspezifisches Unterscheidungsmerkmal darstellt und sich dieses Unterrichtsangebot deshalb nicht nur auf einzelne Schülerinnen und Schüler, sondern in der Regel immer auf ganze Klassen erstreckt (vgl. Abschnitt 1.6.1). Weiterhin konnten nur mithilfe dieses Analyseverfahrens Lehrermerkmale angemessen in die Berechnungen aufgenommen werden, da sich diese immer auf ganze Klasse beziehen und nicht auf die Ebene einzelner Schüler heruntergebrochen werden können. Schließlich wurden mithilfe der eingesetzten Mehrebenenregressionen nicht nur die vergleichenden Leistungsentwicklungsanalysen der Schülergruppen unter gleichzeitiger Konstanzhaltung systematischer Unterschiede zwischen den Klassen (s. oben) möglich – dies stellte das Hauptanliegen von COMBIH dar. Darüber hinaus wurden in der vorliegenden Studie auch die

individuellen Lernvoraussetzungen auf Schülerebene berücksichtigt. Somit konnten zusätzlich Aussagen darüber gemacht werden, welche Schülerinnen und Schüler – unabhängig davon, ob sie Teil eines bilingualen Zugs waren oder nicht – ihre Kompetenzen besonders stark verbesserten; also beispielsweise Mädchen, Lerner mit hohen verbalen Grundfähigkeiten oder Lerner mit einem großen Interesse an Englisch. Folglich leistete die Berücksichtigung dieser individuellen Schülerspekte einen zusätzlichen Beitrag zur Varianzaufklärung der abhängigen Variablen. Eine solche mehrebenenanalytische Auswertung von Daten, die im schulischen Kontext erhoben wurden, stellt eine Standardvorgehensweise der empirischen Bildungsforschung in large-scale Studien dar (z. B. PISA-Studie, PISA-Konsortium, 2001; DESI-Studie, DESI-Konsortium, 2008; vgl. O’Connell & McCoach, 2007) und wurde auch in der vorliegenden Arbeit angewandt. Die hierfür notwendige Unterstützung sowie verschiedene Datenaufbereitungsmaßnahmen (z. B. Birnbaum-Skalierung der Leistungstests; Birnbaum, 1968) erfolgten durch die an der vorliegenden Studie beteiligten empirischen Bildungsforscher.

Diese analytischen Verfahren wurden auch für die Untersuchung eines weiteren Aspekts angewandt, der in der vorliegenden Studie betrachtet wurde: Art und Umfang des Einsatzes von Mutter- und Fremdsprache im bilingualen Unterricht und dessen Wirksamkeit hinsichtlich der Leistungsentwicklung der Klassen in Englisch und Geschichte. Hierbei standen zunächst die Informationen zum Einsatz der Sprachen, welche durch Lehrerfragebögen gewonnen wurden, im Mittelpunkt des Interesses. Deren deskriptive Auswertung ermöglichte Einblicke in eine aktuelle fachdidaktische Fragestellung (vgl. Diehr & Schmelter, 2012) zur Rolle der Sprachen im bilingualen Unterricht. Darüber hinaus wurde der Aspekt der Sprachverwendung in der vorliegenden Studie erstmals in Zusammenhang mit Schülerleistungen gebracht und daraufhin untersucht, wie er diese möglicherweise bedingte. Es konnten also Aussagen dazu getroffen werden, welche sprachbezogenen Unterrichtsmaßnahmen mit günstigen Kompetenzentwicklungen der Klassen einhergingen. Somit beschränkten sich die fachdidaktischen Forschungsanteile der vorliegenden Arbeit nicht nur auf die Erfassung und Analyse von Schülerkompetenzen, wie dies in bisherigen Studien oftmals der Fall war (z. B. Rumlich, 2012; Fehling, 2008). Sie ermöglichten darüber hinaus auch Erkenntnisse zu spezifischen Wirkungsweisen innerhalb des bilingualen Unterrichts. Die hierfür notwendigen Skalen und Items wurden unter Zuhilfenahme der Expertisen von Englisch- und Geschichtsdidaktik entwickelt.

Abschließend soll angemerkt werden, dass eine solche interdisziplinäre Zusammenarbeit mehrerer Wissenschaftsdomänen und die daraus resultierende, multiperspektivische und umfassende Analyse der Wirksamkeit bilingualen Unterrichts jedoch auch zu gewissen Einschränkungen führen kann. Die vorliegende Studie betrachtete beispielsweise lediglich einzelne Aspekte der Englischfähigkeiten und einen nur sehr kleinen spezifischen Ausschnitt aus dem vielfältigen Bereich der Geschichtskompetenzen (vgl. Abschnitt 1.5.2). Überwiegend fachdidaktisch ausgerichtete Studien hingegen könnten, wohl unter Verzicht auf andere in dieser

Studie berücksichtigten Aspekte, den Bereich der Englisch- und vor allem den der Geschichtskompetenzen vertiefen und diese differenzierter und gegebenenfalls auch adäquater mithilfe anderer forschungsmethodischer Zugänge (z. B. offene Aufgabenformate, Unterrichtsbeobachtung) erfassen. Dies trifft auch auf eine umfassendere und detailliertere Betrachtung verschiedenster methodisch-didaktischer Unterrichtsmerkmale des bilingualen Unterrichts zu (vgl. Abschnitt 7.2; sowie Kritik an „Großstudien“ in Brügelmann, 2012).

7.1.3 Implikationen für die Praxis

Indem die Befunde der vorliegenden Arbeit bisherige Forschungsergebnisse zur Wirksamkeit bilingualen Unterrichts mithilfe eines interdisziplinären und analytisch-forschungsmethodisch innovativen Zugangs entscheidend ergänzen konnten (vgl. Abschnitt 7.1.2), besteht deren Bedeutung weiterhin auch in inhaltlichen Implikationen, die sich für die Schul- und Unterrichtspraxis ableiten lassen. Bezüglich der Wirksamkeit bilingualen Unterrichts zeigte sich, dass Schülerinnen und Schüler in der allgemeinen englischen Sprachfähigkeit tendenziell (statistisch jedoch nicht signifikant) sowie hinsichtlich des Hörverstehens stark von dieser Unterrichtsform profitierten. Für die Entwicklung des Sachwissens in Geschichte hingegen konnte kein Effekt des bilingualen Unterrichts festgestellt werden. Diese Befunde können exemplarisch als Basis für zukünftige Entscheidungen bezüglich der Ausweitung bilingualer Unterrichtsprogramme dienen. Die Ergebnisse zu den Hörverstehensleistungen würden dies befürworten, angesichts der großen Vorteile, die sich für die bilingual unterrichteten Klassen ergaben. Allerdings müssten noch weitere Erkenntnisse dazu vorliegen, wie sich diese Unterrichtsform auf andere Fähigkeitsbereiche in Englisch auswirkt, nämlich auf die zweite rezeptive Kompetenz Leseverstehen und die produktiven Fähigkeiten Schreiben und Sprechen. Die Befunde zum Sachwissen in Geschichte könnten gegen eine verstärkte Ausweitung bilingualen Unterrichts in diesem Schulfach sprechen, da sich kein „Mehrwert“ (Hollm, 2013: 185; Breidbach & Viebrock, 2012: 8; Theis, 2010; Barricelli & Schmieder, 2009; Hasberg, 2004) dieser Unterrichtsform für das Fach Geschichte ergab. Darüber hinaus muss beachtet werden, dass die bilingual unterrichteten Klassen der vorliegenden Stichprobe an einer zusätzlichen Geschichtsstunde (d. h. 50% mehr Geschichtsunterricht) pro Woche teilnahmen, im Vergleich zu den einsprachig unterrichteten Schülergruppen (vgl. Teilstudie II). Dies relativiert den Befund vergleichbarer Leistungen zu einem gewissen Grad, da sich bei einer gleichen Anzahl an Unterrichtsstunden vielleicht leichte Nachteile für die bilingual unterrichteten hätten zeigen können. Dennoch kann das Abschneiden der bilingual unterrichteten Lerner in Geschichte auch positiv bewertet werden. Denn auch wenn den Klassen etwas mehr nominale Unterrichtszeit zur Verfügung stand, mussten die Schülerinnen und Schüler zusätzlich zu den Schwierigkeiten, die im Zusammenhang mit dem Erwerb neuen Wissens entstehen, sprachbedingte Herausforderungen meistern und erreichten trotz allem das gleiche Leistungsniveau wie die einsprachig unterrichteten Klassen. Zudem stellt die extra zugewiesene Unterrichtsstunde für den bilingualen

Geschichtsunterricht eine strukturelle Rahmenvorgabe des entsprechenden Ministeriums dar (Ministerium für Kultus, Jugend und Sport Baden-Württemberg), welche faktisch besteht und die deshalb bei Untersuchungen nicht als „Störfaktor“ betrachtet, sondern, im Gegenteil, als Bestandteil des zu untersuchenden Schulsettings gesehen werden sollte. Darüber hinaus deuten die Ergebnisse von Saalbach, Eckstein, Andri, Hobi und Grabner (2013) an, dass sich mithilfe eines leicht veränderten forschungsmethodischen Ansatzes durchaus auch Vorteile für bilingual unterrichtete Klassen im Sachfach ergeben hätten können. Denn die Forschergruppe stellte fest, dass es bei Abweichungen zwischen Unterrichtssprache und Testsprache zu kognitiven Einbußen kommen kann. Dies könnte in der vorliegenden Studie der Fall gewesen sein, da die überwiegende Kommunikationssprache im bilingualen Unterricht zwar Englisch darstellte (vgl. Teilstudie III), der Geschichtstest jedoch vom Großteil der Schülerinnen und Schüler auf Deutsch gelöst werden sollte. Einsprachig unterrichtete Lerner hingegen waren nicht von solchen sprachbezogenen Divergenzen betroffen, weshalb in dieser Schülergruppe „normale“ Leistungen erbracht werden konnten ohne mögliche kognitive Einschränkungen hinnehmen zu müssen. Schließlich konnte die vorliegende Arbeit auch nur Aussagen zu den Effekten bilingualen Unterrichts auf das Sachwissen in Geschichte treffen. Wirksamkeitsanalysen zu anderen Geschichtskompetenzen hingegen, wie etwa Fremdverstehen (Lamsfuß-Schenk, 2008) oder Perspektivenübernahme (Clemen & Sauer, 2007), könnten zu anderen, möglicherweise positiveren Ergebnissen zugunsten der bilingual unterrichteten Klassen führen (vgl. Abschnitt 7.2). In Anlehnung an diese Überlegungen und auf Basis der Befunde der vorliegenden Arbeit zu deutsch-englischem Geschichtsunterricht und den darin berücksichtigten Leistungsbereichen kann zusammenfassend geschlossen werden, dass die verstärkte Etablierung bilingualer Züge an Schulen sinnvoll erscheint. Die Lerner dieses Unterrichtsangebots scheinen im besten Fall bessere Leistungen zu zeigen als ihre einsprachig unterrichteten Mitschülerinnen und Mitschüler. Falls nicht, scheint sich der bilinguale Unterricht jedoch auch nicht negativ auszuwirken und mindestens in vergleichbaren Kompetenzen zu resultieren. Die Überprüfung anderer sachfachlicher Kompetenzbereiche ohne Abweichungen in Kommunikations- und Testsprache dürfte weiterhin zu zusätzlichen positiven Effekten führen. In diesem Sinne kann der in Deutschland nach wie vor vorherrschende Trend zum Ausbau bilingualer Unterrichtsprogramme (KMK, 2013: 10; KM, 2014) auf Grundlage der vorliegenden empirischen Perspektive gestützt werden.

Weiterhin verschafften die vorliegenden Befunde Einblicke in die Wirksamkeit sprachbezogener Unterrichtsmaßnahmen hinsichtlich der Schülerleistung im bilingualen Unterricht. Die entsprechenden Analysen bezogen sich somit lediglich auf das bilinguale Unterrichtssetting und bedurften keines Vergleichs mit einsprachig unterrichteten Klassen. Die Ergebnisse geben Auskunft darüber, welche methodisch-didaktischen Aspekte im Zusammenhang mit der Verwendung von Mutter- und Fremdsprache im bilingualen Unterricht mit günstigeren Kompetenzentwicklungen der Lerner einhergingen und können somit als Orientierung für Lehrkräfte im Unterrichtsgeschehen dienen. Demnach wäre es für bilingual unterrichtende Lehrpersonen empfehlenswert, im bilingualen

Unterricht überwiegend Englisch zu sprechen, da sich dies günstig auf die allgemeine englische Sprachfähigkeit, das Hörverstehen sowie das Sachwissen in Geschichte der Klassen auszuwirken scheint. Dieser Befund bekräftigt entsprechende Empfehlungen, die auf europäischer Ebene (Eurydice, 2006: 8) sowie in der Mehrheit der Bildungs- und Lehrpläne der Bundesländer gegeben wurden (Baden-Württemberg, Bremen, Hamburg, Hessen, Niedersachsen, Rheinland-Pfalz, Sachsen, Thüringen; vgl. Deutscher Bildungsserver, 1996-2014). Darüber hinaus sollte jedoch auch die deutsche Sprache reflektiert im bilingualen Unterricht miteingesetzt werden und zwar insbesondere zur Einführung neuer Fachtermini, da sich dies insbesondere für die rezeptiven Englischfähigkeiten und das Geschichtswissen der Lerner als förderlich erwies. Zudem sollten Lehrkräfte ihre Schülerinnen und Schüler dazu ermutigen, ab und an auch deutsche Redebeiträge abzugeben, da in den vorliegenden Daten ein verstärkter Englischeinsatz seitens der Lerner mit ungünstigeren Kompetenzentwicklungen der Klassen sowohl im Hörverstehen als auch im Geschichtswissen einherging. Diese Befunde zur Verwendung der Muttersprache lassen sich auf die Debatte zur Ein- bzw. Zweisprachigkeit im bilingualen Unterricht beziehen (vgl. Abschnitt 1.4.1) und scheinen den systematischen Miteinbezug des Deutschen nicht nur zu legitimieren, sondern auch einzufordern – eine Position, die sich ebenfalls im fachdidaktischen Diskurs wiederfindet (Heimes, 2013: 349; Diehr, 2012: 25; Gierlinger, Hametner & Spann, 2007: 37; Otten & Wildhage, 2007: 19). Zusammenfassend leisten die vorliegenden Befunde somit einen Beitrag zu der bislang noch weitestgehend ausstehenden (Heimes, 2013: 351; Diehr, 2012: 29) Entwicklung einer empirisch fundierten Methodik bilingualen Unterrichts, die zum Ziel hat, die Kompetenzen der Schülerinnen und Schüler möglichst optimal zu fördern.

Während diese Befunde bisher überwiegend ermutigend erschienen und Impulse für weitere interessante Forschungsfragen zu dieser Unterrichtsform liefern können (s. Folgeabschnitt 7.2), könnten sie jedoch auch Diskussionen zur Gerechtigkeit bilingualer Unterrichtsprogramme anstoßen. Denn die vorliegenden Ergebnisse verweisen, erstens, auf deutliche Unterschiede in den Lernvoraussetzungen zugunsten der bilingual Unterrichteten, welche bereits vor Beginn der Teilnahme am bilingualen Zug bestanden. Zweitens verweisen sie auf teilweise substantielle Vorteile in den Fremdsprachenkompetenzen, die sich mit großer Wahrscheinlichkeit auch aufgrund der Teilnahme an diesem Profil ergaben. Es wäre also zu fragen, ob die zusätzliche Förderung von bereits von Vornhinein begabteren und interessierteren Kindern im Sinne einer „Eliteförderung“ (vgl. Breidbach & Viebrock, 2012: 6; Meyer, 2003) wünschenswert ist und welche ausgleichenden Angebote den Schülerinnen und Schülern außerhalb bilingualer Züge möglicherweise gemacht werden könnten. In diesem Sinne wäre beispielsweise denkbar, diesen Lernern die Teilnahme an zusätzlichem Englischunterricht anzubieten, der in Anlehnung an bilingualen Unterricht besonders kommunikativ und authentisch gestaltet, sowie durch ein geringes Maß an Fehlerkorrektur und einen überwiegenden Verzicht auf Sprachproduktionszwang gekennzeichnet sein müsste. Somit hätten auch die Schülerinnen und Schüler außerhalb bilingualer Unterrichtsprogramme die Chance auf eine

effektive Verbesserung ihrer Fremdsprachenfähigkeiten, ohne jedoch mögliche negative Konsequenzen im Sachfach (z. B. schlechtere Noten) befürchten zu müssen.

Eine weitere Möglichkeit wäre, die durch bilingualen Unterricht verursachten Unterschiede zwischen den Lernern gar nicht erst aufkommen zu lassen, indem das bilinguale Unterrichtsangebot für möglichst viele Schülerinnen und Schüler geöffnet würde. Darauf zielt beispielsweise die Initiative „Bili für alle“ in Nordrhein-Westfalen ab (Breiback & Viebrock, 2012: 6). Entsprechende Empfehlungen finden sich auch in einem Bericht der Kultusministerkonferenz zur aktuellen Situation des bilingualen Unterrichts in Deutschland (KMK, 2013: 20). Diesbezüglich muss jedoch bedacht werden, dass die höheren kognitiven Ansprüche und der gesteigerte Lernaufwand des bilingualen Unterrichts (vgl. Abschnitt 1.2.1) auch einige Lerner überfordern könnten (vgl. Zydati, 2012: 27). Im Sinne eines Kompromisses wäre deshalb denkbar, das bilinguale Unterrichtsangebot in Form von bilingualen Unterrichtsmodulen auszuweiten (vgl. Abschnitt 1.2.1) – ein Trend, der sich in Deutschland bereits beobachten lässt (KMK, 2013: 10). Diesbezüglich müsste jedoch festgestellt werden, ob solche kürzeren und in der Regel unregelmäßiger stattfindenden bilingualen Unterrichtseinheiten dieselbe (positive) Wirkung zeigen wie der Unterricht in fest etablierten bilingualen Zügen. Deshalb scheint eine weitere, in der Schulpraxis bereits umgesetzte und in ihrer Organisationsstruktur vom klassischen Zug-Prinzip abweichende Variante vielversprechend zu sein. Beim sogenannten „Klappklassen“-Prinzip kommen die Lerner, die in das bilinguale Profil einer Schule aufgenommen wurden, immer nur für den bilingualen Unterricht zusammen. In allen anderen Fächern hingegen werden sie in ihren Stammklassen unterrichtet, in denen sich auch Schülerinnen und Schüler anderer (nicht bilingualer) Profile befinden (vgl. Abschnitt 1.2.1). Somit könnten zusätzliche Selektionseffekte zugunsten der bilingual Unterrichteten, die sich – abgesehen von der Auswahl der Lerner zu Beginn des Programms – vor allem im Zuge deren Teilnahme am Programm ergeben, bereits von vornherein abgefangen werden.

7.2 Forschungsperspektiven

Trotz aller bisherigen Analysen besteht noch weiterer Forschungsbedarf zur Wirksamkeit bilingualen Unterrichts, da noch verschiedenste Fragen ungeklärt bleiben. Die vorliegenden Ergebnisse sollten also nicht nur repliziert, sondern vor allem weiter differenziert und ergänzt werden. Dabei sollte die bestmögliche forschungsmethodische Vorgehensweise gewählt werden. Hierzu zählt beispielsweise, erstens, die längsschnittliche Betrachtung der Leistungsentwicklung im bilingualen Unterricht auf, zweitens, der Klassenebene sowie, drittens, die möglichst optimale Isolierung der Effekte bilingualen Unterrichts, um die Wirksamkeit dieser Unterrichtsform nicht zu überschätzen (vgl. Rumlich, 2014a: 83; Abschnitt 1.6.1). Letzteres trifft für zukünftige Untersuchungen insbesondere auf die Englischkompetenzen zu, angesichts der großen Unterschiede, die sich zugunsten der bilingual unterrichteten Klassen ergaben. In diesem Zusammenhang konnten die vorliegenden Befunde bereits zeigen, dass bilingual Unterrichtete zu

Beginn der 8. Klassenstufe über einen Vorteil in Höhe von mindestens einviertel Schuljahren gegenüber den einsprachig Unterrichteten verfügten, obwohl individuelle Lernvoraussetzungen, Motivation und Wissen im Fach Geschichte statistisch konstant gehalten wurden. Darüber hinaus sollte in zukünftigen Studien jedoch auch der Effekt des zusätzlichen vorbereitenden Englischunterrichts berücksichtigt werden, welchen Lerner bilingualer Züge oft bereits vor Beginn des zweisprachigen Fachunterrichts in den Klassenstufen 5 und 6 erteilt bekommen (vgl. Abschnitt 1.2.1). Solche Untersuchungen würden Datenerhebungen schon zu Beginn der 5. und am Ende der 6. Klassenstufe, also zum Anfang und Ende dieses vorbereitenden Englischunterrichts, erfordern (vgl. Bos, Bonsen & Gröhlich, 2009). Weitere Messzeitpunkte könnten dann ab Klassenstufe 7, jeweils zu Beginn und Ende des Schuljahres, in den jeweiligen bilingualen Unterrichtsfächern durchgeführt werden (vgl. Rumlich, 2012; Fehling, 2008). Abgesehen vom verstärkten Fremdsprachenunterricht in den Klassen 5 und 6 müssten zukünftig jedoch noch weitere Faktoren erfasst werden, die zum Leistungsvorsprung der bilingual unterrichteten Schülerinnen und Schüler in Englisch beitragen könnten. Diesbezüglich wären beispielsweise außerschulische Kontakte der Lerner zu dieser Fremdsprache von Interesse, zum Beispiel in Form von Emailbekanntschaften und Besuch von Verwandten oder regelmäßigen Urlauben im englischsprachigen Ausland (vgl. Köller et al., 2012: 348; Rumlich, 2012: 171), sowie insbesondere auch Einstellung und Kontakte der Eltern zur Fremdsprache (Helmke, 2012). Diese Vorgehensweise würde zu besonders belastbaren Befunden zur Wirksamkeit bilingualen Unterrichts in Bezug auf die Englischkompetenzen der Lerner führen, da solche Ergebnisse bereinigt wären von vielen diese Leistung entscheidend mitbeeinflussenden Aspekten.

Weiterhin erfordern möglichst adäquate Aussagen zur Wirksamkeit dieser Unterrichtsform nicht nur die statistische Kontrolle von (außerschulischen) Faktoren, die die Englischkompetenzen der Schülerinnen und Schüler meist schon vor Beginn des bilingualen Unterrichts bedingen, sondern auch solcher, die Schülerleistung generell und insbesondere auch während des bilingualen Unterrichts beeinflussen (vgl. Pérez-Cañado, 2012: 330). In Bezug auf Helmkes (2012) Angebots-Nutzungs-Modell zur Wirkungsweise des Unterrichts betrifft dies mehrere Bereiche. Ein erster solcher Bereich stellen die individuellen Lernvoraussetzungen der Schülerinnen und Schüler dar, die in zukünftigen Untersuchungen erfasst werden sollten – wie in der vorliegenden Arbeit und anderen Studien bereits geschehen (Rumlich, 2014b; Köller et al., 2012). Weiterhin könnte dies um weitere Aspekte ergänzt werden, wie beispielweise um Persönlichkeitsmerkmale der Lerner (z. B. Gewissenhaftigkeit aus den „Big Five“; Lang & Lüdtke, 2005). Denn diesbezüglich kann angenommen werden, dass bilingual unterrichtete Schülerinnen und Schüler – möglicherweise aufgrund ihrer größeren Motivation und ihrem gesteigerten Interesse an Schule und Unterricht – auch gewissenhafter sein könnten als ihre einsprachig unterrichteten Mitschülerinnen und Mitschüler, was sich wiederum positiv auf deren Schulleistungen auswirken kann (Poropat, 2009).

Ein weiterer Aspekt, der im Zusammenhang mit den Lernvoraussetzungen steht und die Schülerleistungen ebenfalls beeinflusst, stellt die Klassenzusammensetzung dar. Zukünftige Studien könnten diesbezüglich nicht nur potentielle Unterschiede zwischen den Klassen kontrollieren, wie in der vorliegenden Studie geschehen, sondern darüber hinaus Kompositionseffekte in den Blick nehmen (z. B. Dumont, Neumann, Nagy, Becker, Rose & Trautwein, 2013; Neumann, Milek, Maaz & Gresch, 2010). Hierdurch könnte beispielsweise die Frage beantwortet werden, ob es unabhängig von der individuellen Leistung eines Lerners einen Unterschied macht, dass sich dieser in einer leistungsstarken oder leistungsschwachen Klasse befindet. Die Berücksichtigung dieser Kompositionsfragestellung scheint im Kontext des bilingualen Unterrichts von besonderer Relevanz, da die vorliegenden Befunde auf ein deutliches Gefälle im Leistungsniveau zwischen bilingual und einsprachig unterrichteten Klassen zugunsten der bilingual unterrichteten verwiesen.

Ein dritter Bereich, der nach Helmke (2012) die Schülerleistungen beeinflusst und deshalb in zukünftigen Studien ebenfalls in den Blick genommen werden sollte, stellt die individuelle Wahrnehmung des bilingualen Unterrichtsangebots durch die Schülerinnen und Schüler dar. Dies scheint im vorliegenden Kontext deshalb besonders relevant, da angenommen werden kann, dass bilingual unterrichtete den Unterricht positiver wahrnehmen als einsprachig unterrichtete Lerner (vgl. Abschnitt 1.4.4). Zur Untersuchung dieser Fragestellung könnten die in der vorliegenden Arbeit erhobenen Daten zu Leistungsemotionen analysiert sowie weitere individuelle Interpretationsmuster der Lerner erfasst werden. Dasselbe trifft auf die individuelle Nutzung des Unterrichtsangebots durch die Schülerinnen und Schüler zu. Diesbezüglich könnten zukünftige Studien beispielsweise schülerspezifische Denkprozesse, schulische und außerschulische Lernaktivitäten oder die aktiv genutzte Lernzeit im Unterricht erheben (vgl. Helmke, 2012) und diese in Analysen im Zusammenhang mit der Kompetenzentwicklung der Schülerinnen und Schüler im bilingualen Unterricht berücksichtigen. Da hierzu noch keine Forschungsarbeiten oder sonstigen Erfahrungswerte bekannt sind, bestehen bisher keine Vermutungen dazu, ob sich bilingual von einsprachig unterrichteten Lernern in diesen Aspekten unterscheiden könnten. Dies stellt daher eine offene Forschungsfrage dar.

Abschließend sollte in künftigen Untersuchungen zur Wirksamkeit bilingualen Unterrichts unbedingt auch die Lehrperson Berücksichtigung finden (vgl. Helmke, 2012). Hierfür könnten die in der vorliegenden Arbeit erhobenen Daten zu bilingual und einsprachig unterrichtenden Lehrkräften daraufhin analysiert werden, ob sich systematische Unterschiede zugunsten der bilingual unterrichtenden Kolleginnen und Kollegen ergeben. Dies könnte beispielsweise bezüglich deren schul- und unterrichtsbezogenen Einstellungen, Erwartungen und Ziele oder hinsichtlich ihres unterrichtlichen Engagements der Fall sein (Hollm, 2013: 185; Dalton-Puffer, Nikula & Smit, 2010: 282; vgl. Abschnitt 1.3.1), wie ansatzweise bereits gezeigt in der vorliegenden Studie (vgl. Teilstudie II). Darüber hinaus sollte jedoch auch das Professionswissen der Lehrerinnen und Lehrer, insbesondere in fachwissenschaftlicher und fachdidaktischer Hinsicht (vgl. Kunter, Baumert, Blum,

Klusmann, Krauss & Neubrand, 2011), sowie deren pädagogische Orientierungen umfassend erhoben und in Untersuchungen zur Kompetenzentwicklung bilingual und einsprachig unterrichteter Klassen statistisch konstant gehalten werden. Denn wie bereits in Abschnitt 1.3.3 berichtet, beeinflusst die Lehrperson und ihr Unterricht den Schülerertrag nicht unwesentlich mit (Helmke, 2012; Hattie, 2009; Klieme & Rakoczy, 2008). Dies trifft ebenfalls auf allgemeine Unterrichtsqualitätsmerkmale zu. Die diesbezüglich in der vorliegenden Arbeit berücksichtigten Dimensionen (Klassenführung, konstruktive Unterstützung, kognitive Aktivierung) könnten um weitere Aspekte erweitert werden, wie etwa Methodenvielfalt/-abwechslung, Umgang mit Heterogenität und individuelle Förderung der Schülerinnen und Schüler (Meyer, 2005) sowie Klassenklima und zur Verfügung stehende Unterrichtszeit (Helmke, 2012). Insbesondere bezüglich dieses letzten Punkts wäre es für künftige Studien interessant, Wirksamkeitsanalysen zu bilinguaem Unterricht an einer Stichprobe durchzuführen, die dieselbe wöchentliche Stundenzahl an Sachfachunterricht für bilingual und einsprachig unterrichtete Klassen aufweist (vgl. Abschnitt 1.2.1). Solche Untersuchungen könnten zeigen, ob die Leistungsentwicklung beider Schülergruppen im Sachfach auch unter diesen Bedingungen noch vergleichbar bliebe oder vielleicht etwas ungünstiger für die bilingual Unterrichteten verlaufen würde. Andererseits könnten sich im Sachfach auch leichte Vorteile für bilingual unterrichtete Klassen zeigen, wenn in zukünftigen Studien die sachfachlichen Leistungen in derselben Sprache geprüft würden, welche zur Kommunikation im Unterricht verwandt wurde (Saalbach et al., 2013; vgl. Abschnitt 7.1.3).

Die bisher aufgezeigten Forschungsdesiderate sind bei Untersuchungen zu Effekten bilingualen Unterrichts auf die Schülerleistung nicht nur aus inhaltlichen, sondern überwiegend auch aus forschungsmethodischen Gründen von Relevanz und sollten deshalb in entsprechenden zukünftigen Studien im Idealfall in den Blick genommen werden. Darüber hinaus kann die Wirksamkeit dieser Unterrichtsform in Ergänzung zu den Befunden der vorliegenden Arbeit auch über andere Aspekte und in anderen Kontexten bestimmt werden, da sich bilingualer Sachfachunterricht nicht nur auf die Fächer Englisch und Geschichte an Gymnasien bezieht (vgl. KMK, 2013: 11-14) und sich dabei auch nicht nur auf die fachspezifischen Kompetenzen der Schülerinnen und Schüler auswirkt. In diesem Sinne sollten in zukünftigen Studien zur Wirksamkeit dieser Unterrichtsform weitere Zielkriterien (Helmke, 2012: 82) untersucht werden. Dies bezieht sich zunächst auf eine differenziertere Erfassung der Englischleistungen, die, abgesehen von der allgemeinen Sprachfähigkeit und dem Hörverstehen der Lerner, auch in Bezug auf die zweite rezeptive Fähigkeit, das Leseverstehen, sowie in Bezug auf produktive Fertigkeiten (Schreiben, Sprechen) und andere Kompetenzbereiche (z. B. interkulturelle Kompetenz) erfolgen sollte (vgl. KMK, 2004: 8). Bisherige Untersuchungen berichteten beispielsweise uneinheitliche Ergebnisse bezüglich der Schreibkompetenz bilingual Unterrichteter. Sie verwiesen entweder auf einen positiven (Ruiz de Zarobe, 2010; Zydariß, 2007) oder auf keinen nennenswerten Effekt (Nold et al., 2008) bilingualen Unterrichts. Weiterhin deuten erste Studien Vorteile bilingual Unterrichteter im

Bereich Fremdsprachenvokabular an (Ruiz de Zarobe, 2010; Sylvén, 2010), was in zukünftigen Studien validiert werden müsste.

Ähnliches gilt für die Leistungen von Schülerinnen und Schülern im Fach Geschichte. Diesbezüglich sollten nicht nur das Sachwissen der Lerner Berücksichtigung finden und die Daten der vorliegenden Arbeit zu deren Fähigkeit, den Konstruktionscharakter von Geschichte zu erkennen, analysiert werden. Darüber hinaus steht eine Vielzahl an weiteren Teilkompetenzen aus verschiedenen geschichtsdidaktischen Modellen (z. B. Gautschi, 2009; Pandel, 2005) sowie die ihnen zugrunde liegende Hauptdimension „Geschichtsbewusstsein“ zur Verfügung (vgl. Abschnitt 1.5.2), die in zukünftigen Analysen zur Wirksamkeit bilingualen Unterrichts hinsichtlich der Leistungen in Geschichte berücksichtigt werden könnten. Hierbei wären insbesondere die Kompetenzen Fremdverstehen, Multiperspektivität und die Fähigkeit zum Perspektivenwechsel zu untersuchen, da verschiedene Längsschnittstudien diesbezüglich bereits positive Ergebnisse zugunsten der bilingual unterrichteten berichteten (Lamsfuß-Schenk, 2008; Clemen & Sauer, 2007). Die Befunde wurden von den Autoren auf die bessere Tiefenverarbeitung und Durchdringung der Inhalte durch die bilingual unterrichteten Lerner zurückgeführt, was wiederum mit deren längeren und sorgfältigeren Auseinandersetzung mit den fremdsprachigen (und deshalb anspruchsvolleren) Materialien erklärt wurde. Diese Annahmen werden von weiteren querschnittlichen Studien gestützt, deren Ergebnisse auf ein besseres Textverständnis bilingual Unterrichteter und, in Folge dessen, auf eine detailliertere, differenziertere, ausführlichere und kohärentere Wiedergabe dieser Inhalte sowie auf eine präzisere Verwendung von Fachbegriffen verwiesen (Bernhardt, 2011; Gruner, 2011; Madrid, 2011; Lamsfuß-Schenk, 2008; Clemen & Sauer, 2007). Um den Anforderungen dieser Kompetenzbereiche gerecht zu werden, bieten sich zu deren Untersuchung neben den in der vorliegenden Studie gewählten Instrumenten, auch andere Erfassungsmethoden an, wie beispielsweise Tests mit offenen Aufgabenformaten oder Testinterviews (vgl. Abschnitt 7.1.2).

Weitere Zielkriterien, anhand derer die Wirksamkeit bilingualen Unterrichts festgestellt werden kann, sind die Entwicklung von Motivation und Emotionen der Schülerinnen und Schüler in dieser Unterrichtsform. In diesem Sinne könnte mithilfe der in der vorliegenden Arbeit erhobenen, jedoch noch nicht ausgewerteten, Daten untersucht werden, ob sich diese Entwicklungsverläufe bei den bilingual unterrichteten Klassen möglicherweise positiver gestalten als bei den einsprachig unterrichteten Schülergruppen. Hierzu bestehen bereits erste Überlegungen und empirische Befunde (Piesche, Jonkmann & Keßler, 2014; Rumlich, 2014a; Abendroth-Timmer, 2007). Diese weisen jedoch auf Widersprüchlichkeiten hin und müssten somit in künftigen Studien weiter validiert werden. Schließlich wäre es auch lohnenswert, den Aspekt des Empathievermögens bzw. der emotionalen Intelligenz als Teil der sozialen Kompetenz (Mayer & Salovey, 1990: 190) im Kontext der Wirksamkeit bilingualen (Geschichts-)Unterrichts zu betrachten (Heimes, 2013: 347). Denn diesbezüglich kann angenommen werden, dass bilingual unterrichtete Lerner aufgrund der Verwendung zweier Unterrichtssprachen Fremdverstehen und Multiperspektivität (s. oben) besser

ausbilden könnten als einsprachig unterrichtete Klassen (vgl. Geiss, 2013; Heimes, 2013: 347) und sich somit besser in die Lage Anderer versetzen und dementsprechend angemessener interagieren können. Abschließend muss jedoch darauf hingewiesen werden, dass sich bei der Betrachtung einzelner Zielkriterien im Kontext von Wirksamkeitsanalysen *aptitude-treatment-interactions* (Snow, 1989) ergeben können. Demnach muss in Forschungsarbeiten mitbedacht werden, dass das Erreichen eines Ziels (z. B. reine Leistungssteigerung) zu negativen Effekten bezüglich eines anderen Zielkriteriums (z. B. Ausbildung der Motivation) führen kann (Helmke & Weinert, 1997: 150). Dies trifft sowohl auf den bilingualen als auch auf jeden anderen Unterricht zu. Angesichts dieser Problematik könnten zukünftige Untersuchungen bei der Analyse von Daten auf Strukturgleichungsmodelle zurückgreifen, die solche Effekte abbilden können (Christ & Schlüter, 2011).

Abgesehen vom Aspekt des Zielkriteriums bestehen weiterhin verschiedenste Kontexte, innerhalb derer die Wirksamkeit bilingualen Unterrichts untersucht werden kann. Dies bezieht sich zum einen auf die im bilingualen Unterricht verwandte Fremdsprache. Hierfür wären weitere Studien zu deutsch-französischem Sachfachunterricht sowie zu „exotischeren“ Sprachkombinationen wie beispielsweise Deutsch–Italienisch oder Deutsch–Polnisch wünschenswert (vgl. KMK, 2013: 13). Zum anderen sollten sich zukünftige Studien nicht nur auf den bisher dominantesten Bereich der geisteswissenschaftlichen Schulfächer, wie Geschichte, Geographie oder Politik, begrenzen (Pérez-Cañado, 2012: 320; Mentz, 2010: 33), sondern auch Wirksamkeitsanalysen in Mathematik, in naturwissenschaftlichen Fächern, in Musik oder in Sport durchführen. Dies ist nicht zuletzt deshalb von Interesse, da in Deutschland mittlerweile in fast jedem Schulfach bilinguale Unterrichtsangebote gemacht werden (KMK, 2013: 14), zu deren Wirksamkeit in Bezug auf die Sachfachkompetenzen oder anderen Zielkriterien jedoch kaum Befunde vorliegen (Bonnet, 2012: 67; Bernhardt, 2011: 215; Coyle et al., 2010: 138; Sylvén, 2010: 224). Ähnliches gilt für die Schulform. Bisher nahm die Mehrheit der Studien zum bilingualen Unterricht dieses spezielle Unterrichtsangebot lediglich im Kontext des Gymnasiums in den Blick. In den letzten Jahren wurden jedoch – vermutlich angesichts der verstärkten Ausweitung bilingualer Unterrichtsangebote (KMK, 2013) – auch vermehrt Studien an Realschulen (z. B. Hartmannsgruber, 2014; Piesche et al., 2014; Hollm, Hüttermann, Keßler, Schlemminger & Ade-Thurow, 2013) und Grundschulen (z. B. Steinlen & Piske, 2015; Merisuo-Storm & Soininen, 2014; Haudeck, McCafferty & Aristov, 2012; Massler, 2010) sowie einige wenige Forschungsvorhaben auch an Hauptschulen (z. B. Schwab, 2013) durchgeführt. Diese Untersuchungen innerhalb verschiedener Schulformen sollten zukünftig weiter vertieft werden. Damit in Zusammenhang steht auch der Bedarf an weiteren Analysen in verschiedenen Altersstufen. Darüber hinaus wäre insbesondere in Baden-Württemberg der Forschungskontext der im Schuljahr 2012/13 neu eingeführten Gemeinschaftsschulen von Interesse (Landesbildungsserver Baden-Württemberg, 2014). Aufgrund der dort erwarteten heterogenen Schülerschaft, die möglicherweise besondere Ansprüche an den

bilingualen Unterricht stellt, wären Analysen zur Implementation und Wirksamkeit dieser Unterrichtsform in dieser Schulform lohnenswert. Schließlich müsste die Wirksamkeit bilingualen Unterrichts auch über den Schulkontext hinaus in den Blick genommen werden. Hierbei wären insbesondere Längsschnittstudien interessant, die ehemals bilingual sowie einsprachig unterrichtete Schülerinnen und Schüler in ihrem weiteren Ausbildungsweg und in das Berufsleben hinein begleiten und diesbezüglich mögliche Unterschiede analysieren. Auf Grundlage der momentanen Attraktivität und globalen Relevanz von Englisch als *lingua franca* sowie angesichts der festgestellten Stärken bilingual Unterrichteter in diesem Bereich, könnte angenommen werden, dass diese – zumindest in den Anfangsjahren nach der Schulzeit – erfolgreichere Karriereschritte absolvieren als ehemals einsprachig unterrichtete Schülerinnen und Schüler.

Überdies sollten im Mittelpunkt zukünftiger Studien verstärkt mögliche Wirkungsweisen spezifischer methodisch-didaktischer Merkmale bilingualen Unterrichts stehen, indem deren Effekte auf verschiedene Zielkriterien, wie beispielsweise Kompetenzentwicklung, untersucht werden. So könnten ausgehend von den Ergebnissen der vorliegenden Arbeit zu Art und Umfang des Einsatzes von Mutter- und Fremdsprache weitere Unterrichtsaspekte betrachtet werden, wie beispielsweise der Einsatz von Sozialformen. Hierzu liegen bereits erste deskriptive Befunde vor (z. B. Kollenrott, 2008; Viebrock, 2007; vgl. Abschnitt 1.6.4), die nun noch in Bezug zu Schülerleistungen gesetzt werden müssten. Weitere Untersuchungen könnten sich mit der Qualität und den Prinzipien der Gestaltung von Unterrichtsmaterial (z. B. didaktische Reduktion, Sprachverwendung, Visualisierung etc.) beschäftigen sowie mit dem Einsatz von Medien, mit Unterstützungssystemen seitens der Lehrkräfte oder Art und Häufigkeit der Korrektur von Schülerfehlern. Hierfür würden sich neben Befragungen auch andere Untersuchungsmethoden anbieten, wie etwa die Unterrichtsbeobachtung oder die Videographie, welche die unterrichtlichen Geschehnisse gegebenenfalls noch präziser erfassen können. Entsprechende Befunde würden wichtige Impulse für eine empirisch fundierte Didaktik und Methodik dieser Unterrichtsform liefern. Dies scheint insbesondere deshalb relevant, da in den vergangenen Jahren bereits des Öfteren festgestellt wurde, dass noch keine einheitliche Didaktik und Methodik für den bilingualen Sachfachunterricht existiert (KMK, 2013: 7; Thürmann, 2013: 230; Rautenhaus, 2010: 109; Breidbach, 2010: 174).

Zusammenfassend verweisen diese Überlegungen auf eine Vielzahl an Aspekten und Kontexten, die im Hinblick auf die Wirksamkeit und spezifische Wirkungsweise des bilingualen Unterrichts in zukünftigen Studien noch untersucht werden könnten. Hierfür sollte angesichts der inhaltlichen und forschungsmethodisch bedingten Herausforderungen eines solch komplexen Forschungsgegenstands eine interdisziplinäre, sich in ihren jeweiligen Stärken ergänzende, Zusammenarbeit zwischen Sprach- und Sachfachdidaktiken sowie der empirischen Bildungsforschung angestrebt werden, um die umfassende und möglichst adäquate Erfassung dieser speziellen Unterrichtsform zu gewährleisten.

7.3 Fazit

Die Bedeutung der vorliegenden Arbeit zur Wirksamkeit deutsch-englischen Geschichtsunterrichts liegt darin, dass sie die bisherigen Befunde zu diesem Forschungsfeld mithilfe eines analytisch-forschungsmethodisch innovativen Zugangs replizieren und darüber hinaus entscheidend differenzieren und ergänzen konnte. Dies wurde in Anlehnung an Helmkes (2012) Angebots-Nutzungs-Modell durch die interdisziplinäre Zusammenarbeit zwischen Englisch- und Geschichtsdidaktik sowie der empirischen Bildungsforschung ermöglicht. Dadurch konnten Fragestellungen aus allen drei Wissenschaftsdisziplinen berücksichtigt, angemessene forschungsmethodisch-analytische Verfahren angewandt und somit die Wirksamkeit dieser Unterrichtsform möglichst umfassend und adäquat untersucht werden. Die vorliegenden Befunde können einerseits als Basis für zukünftige Entscheidungen hinsichtlich der Ausweitung bilingualer Unterrichtsprogramme dienen. Andererseits leisten sie einen Beitrag zu einer empirisch fundierten Methodik des bilingualen Sachfachunterrichts. Zukünftige Studien könnten die Wirksamkeit dieser Unterrichtsform hinsichtlich anderer Zielkriterien, beispielsweise Motivation oder Emotionen der Lerner, sowie innerhalb anderer Kontexte, wie etwa in anderen Schulfächern, Schulformen oder im Berufsleben, bestimmen und darüber hinaus weitere spezifische methodisch-didaktische Merkmale bilingualen Unterrichts und deren Wirkungsweise in den Blick nehmen.

7.4 Literatur

- Abendroth-Timmer, Dagmar (2007). *Akzeptanz und Motivation: Empirische Ansätze zur Erforschung des unterrichtlichen Einsatzes von bilingualen und mehrsprachigen Modulen*. Frankfurt/ Main: Lang.
- Adesope, Olusola. O.; Lavin, Tracey; Thompson, Terri & Ungerleider, Charles (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research* 80(2), 207–245.
- Barricelli, Michele & Schmieder, Ulrich (2009). Über Nutzen und Nachteil des bilingualen Sachfachunterrichts. Fremdsprachen- und Geschichtsdidaktik im Dialog. In Daniela Caspari, Wolfgang Hallet, Anke Wegner & Wolfgang Zydati (Hrsg.), *Bilingualer Unterricht macht Schule. Beitrge aus der Praxisforschung*. Frankfurt/ Main: Lang, 205–220.
- Bernhardt, Markus (2011). Bilingualitt und historisches Lernen. Frderung von historischen Kompetenzen oder soziales Differenzkriterium? In Jan Hodel & Batrice Ziegler (Hrsg.), *Forschungswerkstatt Geschichtsdidaktik 09. Beitrge zur Tagung "Geschichtsdidaktik Empirisch 09" vom 3. und 4. September 2009 in Basel*. Bern: hep, 214–223.
- Birnbaum, Allan (1968). Some latent trait models and their use in inferring an examinee's ability. In Frederic M. Lord & Melvin R. Novick (Hrsg.), *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley, 397–479.
- Bonnet, Andreas (2012). Towards an evidence base for CLIL: How to integrate qualitative and quantitative as well as process, product and participant perspectives in CLIL research. *International CLIL Research Journal* 4(1), 66–78. URL: <http://www.icrj.eu/14/article7.html>, Zugriff am 08.12.2014.
- Bos, Wilfried; Bosen, Martin & Grhlich, Carola (2009). *KESS 7: Kompetenzen und Einstellungen von Schlerinnen und Schlern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*. Mnster: Waxmann.
- Breidbach, Stephan (2010). Bilinguale Didaktik – bald wieder zwischen allen Sthlen? Zu den Aussichten einer integrativen Didaktik des biligualen Sachfachunterrichts. In Gerhard Bach & Susanne Niemeier (Hrsg.), *Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven*. Frankfurt/ Main: Lang, 165–176.
- Breidbach, Stephan & Viebrock, Britta (2012). CLIL in Germany: Results from recent research in a contested field of education. *International CLIL Research Journal* 4(1), 5–16. URL: <http://www.icrj.eu/14/article1.html>, Zugriff am 08.12.2014.
- Brgelmann, Hans (2012). Begrenzter Erkenntniswert. PISA Kontra. In *Erziehung und Wissenschaft* 64(7-8), 19.
- Butzkamm, Wolfgang (1973). *Aufgeklrte Einsprachigkeit. Zur Entdogmatisierung der Methode im Fremdsprachenunterricht*. Heidelberg: Quelle & Meyer.

- Christ, Oliver & Schlüter, Elmar (2011). *Strukturgleichungsmodelle mit Mplus: Eine praktische Einführung*. München: Oldenbourg.
- Clemen, Franziska & Sauer, Michael (2007). Förderung von Perspektivendifferenzierung und Perspektivenübernahme? Bilingualer Geschichtsunterricht und historisches Lernen. Eine empirische Studie. *Geschichte in Wissenschaft und Unterricht* 58(12), 708–723.
- Coyle, Do; Hood, Philip & Marsh, David (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Cummins, Jim (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49(2), 222–251.
- Dalton-Puffer, Christiane; Nikula, Tarja & Smit, Ute (2010). Language use and language learning in CLIL. Current findings and contentious issues. In Christiane Dalton-Puffer, Tarja Nikula & Ute Smit (Hrsg.), *Language use and language learning in CLIL classrooms*. Amsterdam: Benjamins, 279–291.
- DESI-Konsortium (Hrsg.) (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Deutscher Bildungsserver (1996 – 2014). *Bildungspläne / Lehrpläne der Bundesländer für allgemeinbildende Schulen*. URL: <http://www.bildungsserver.de/Bildungsplaene-Lehrplaene-der-Bundeslaender-fuer-allgemeinbildende-Schulen-400.html>, Zugriff am 26.09.2014.
- Diehr, Bärbel (2012). What's in a name? Terminologische, typologische und programmatische Überlegungen zum Verhältnis der Sprachen im bilingualen Unterricht. In Bärbel Diehr & Lars Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang, 17–36.
- Diehr, Bärbel & Schmelter, Lars (Hrsg.) (2012). *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang.
- Dumont, Hanna; Neumann, Marko; Nagy, Gabriel; Becker, Michael; Rose, Norman & Trautwein, Ulrich (2013). Einfluss der Klassenkomposition auf die Leistungsentwicklung in Haupt- und Realschulen in Baden-Württemberg. *Psychologie in Erziehung und Unterricht* 60, 198–213.
- Eurydice (Hrsg.) (2006). *Content and Language Integrated Learning (CLIL) at school in Europe*. URL: http://www.indire.it/lucabas/lkmw_file/eurydice/CLIL_EN.pdf, Zugriff am 09.12.2014.
- Fehling, Sylvia (2008). *Language Awareness und bilingualer Unterricht: Eine komparative Studie*. Frankfurt/ Main: Lang.
- Gass, Susan M. & Mackey, Alison (2007). Input, interaction, and output in second language acquisition. In Bill VanPatten & Jessica Williams (Hrsg.), *Theories in second language acquisition. An introduction*. Mahwah, NJ: Erlbaum, 175–199.

- Gautschi, Peter (2009). *Guter Geschichtsunterricht: Grundlagen, Erkenntnisse, Hinweise*. Schwalbach/ Ts.: Wochenschau Verlag.
- Geiss, Peter (2013). Bilingualer Geschichtsunterricht. Ein Modell für das historische Lernen im *global village*. In Bärbel Diehr & Lars Schmelter (Hrsg.), *Bilingualen Unterricht weiterdenken. Programme, Positionen, Perspektiven*. Frankfurt/ Main: Lang, 55–71.
- Gierlinger, Erwin; Hametner, Christian & Spann, Harald (2007). Englisch als Arbeitssprache – viele Inseln und (noch) kein Festland? Eine Studie zum Arbeitssprachenunterricht auf Sekundarstufe I in Oberösterreich. In: Erwin Gierlinger, Verena Hainschink & Harald Spann (Hrsg.), *Content and Language Integrated Learning (CLIL) auf der Sekundarstufe I. Entwicklung, Forschung und Praxis*. Linz: Trauner, 63–118.
- Gruner, Carola (2011). Entwicklung und Förderung historischer Kompetenzen im bilingualen Geschichtsunterricht. In Jan Hodel & Béatrice Ziegler (Hrsg.), *Forschungswerkstatt Geschichtsdidaktik 09. Beiträge zur Tagung „Geschichtsdidaktik empirisch“*. Bern: hep-Verlag, 202–213.
- Hartmannsgruber, Martin (2014). *Bilinguale Biologie: Konzeption und Evaluation*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Hasberg, Wolfgang (2004). Historisches Lernen im bilingualen Geschichtsunterricht(?). In Andreas Bonnet & Stephan Breidbach (Hrsg.), *Didaktiken im Dialog. Konzepte des Lehrens und Wege des Lernens im bilingualen Sachfachunterricht*. Frankfurt/ Main: Lang, 221–236.
- Hattie, John (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Haudeck, Helga; McCafferty, Susanne & Aristov, Natasha (2012). Bilingualer Sachfachunterricht in der Grundschule am Beispiel von Storytelling und Experimenten zu Naturphänomenen. In Heiner Böttger & Norbert Schlüter (Hrsg.), *Fortschritte im Frühen Fremdsprachenlernen*. München: Domino Verlag, 157–165.
- Heimes, Alexander (2013). Didaktiken und Methodiken bilingualer Fächer: Geschichte. In Wolfgang Hallet & Frank G. Königs (Hrsg.), *Handbuch Bilingualer Unterricht. Content and Language Integrated Learning*. Seelze: Kallmeyer, 345–352.
- Helmke, Andreas (2012). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Kallmeyer.
- Helmke, Andreas & Weinert, Franz E. (1997). Bedingungsfaktoren schulischer Leistung. In Franz E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule. Enzyklopädie der Psychologie: Pädagogische Psychologie*. Göttingen: Hogrefe, 71–176.
- Hollm, Jan (2013). Bilinguales Lehren und Lernen als schulische Innovation aus Sicht der Beteiligten. In Jan Hollm, Armin Hüttermann, Jörg-U. Keßler, Gérald Schlemminger & Benjamin Ade-Thurow (Hrsg.), *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik, 167–192.

- Hollm, Jan; Hüttermann, Armin; Keßler, Jörg-U.; Schlemminger, Gérald & Ade-Thurow, Benjamin (Hrsg.) (2013). *Bilinguales Lehren und Lernen in der Sekundarstufe I: Sprache, Sachfach und Schulorganisation*. Landau: Verlag Empirische Pädagogik.
- Klieme, Eckard & Rakoczy, Katrin (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik* 54(2), 222–237.
- KM (Ministerium für Kultus, Jugend und Sport Baden-Württemberg) (Hrsg.) (2014). *Ausweitung bilingualer Züge an Realschulen*. URL: <http://www.kultusportal-bw.de/Lde/770313?QUERYSTRING=bilingual>, Zugriff am 26.09.14.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Hrsg.) (2013). *Bericht „Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung“: Beschluss der Kultusministerkonferenz vom 17.10.2013*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/201_10_17-Konzepte-_bilingualer-_Unterricht.pdf, Zugriff am 09.12.2014.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Hrsg.) (2004). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-erste-Fremdsprache.pdf, Zugriff am 09.12.2014.
- Kollenrott, Anne (2008). *Sichtweisen auf deutsch-englisch bilingualen Geschichtsunterricht: Eine empirische Studie mit Fokus auf interkulturelles Lernen*. Frankfurt/ Main: Lang.
- Kölller, Olaf; Leucht, Michael & Pant, Hans (2012). Effekte bilingualen Unterrichts auf die Englischleistungen in der Sekundarstufe I. *Unterrichtswissenschaft* 4(4), 334–350.
- Krashen, Stephen D. (1985). *The input hypothesis. Issues and implications*. London: Longman.
- Krashen, Stephen D. & Terrell, Tracey D. (2000). *The natural approach. Language acquisition in the classroom*. Harlow: Pearson Education.
- Kunter, Mareike; Baumert, Jürgen; Blum, Werner; Klusmann, Uta; Krauss, Stefan & Neubrand, Michael (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Lamsfuß-Schenk, Stefanie (2008). *Fremdverstehen im bilingualen Geschichtsunterricht. Eine Fallstudie*. Frankfurt/ Main: Lang.
- Landesbildungsserver Baden-Württemberg (2014). *Gemeinschaftsschule*. URL: <http://www.schule-bw.de/schularten/gemeinschaftsschule/>, Zugriff am 09.12.2014.
- Lang, Frieder R. & Lütke, Oliver (2005). Der Big Five-Ansatz der Persönlichkeitsforschung: Instrumente und Vorgehen. In Siegfried Schumann (Hrsg.), *Persönlichkeit: eine vergessene Größe der empirischen Sozialforschung*. Wiesbaden: VS, 29–39.

- Madrid, Daniel (2011). Monolingual and bilingual students' competence in social sciences. In Daniel Madrid & Stephen Hughes (Hrsg.), *Studies in bilingual education*. Bern: Lang, 195–222.
- Massler, Ute (2010). Schwierig aber machbar: Leistungsbewertung im CLIL-Unterricht der Grundschule. In Ute Massler & Petra Burmeister (Hrsg.), *CLIL und Immersion: Fremdsprachlicher Sachfachunterricht in der Grundschule*. Braunschweig: Westermann, 131–142.
- Mayer, John D. & Salovey, Peter (1990). Emotional intelligence. *Imagination, Cognition, and Personality* 9, 185–211.
- Mentz, Oliver (2010). Alle Fächer eignen sich – oder doch nicht? Überlegungen zu einem bilingualen Fächerkanon. In Sabine Doff (Hrsg.), *Bilingualer Sachfachunterricht in der Sekundarstufe. Eine Einführung*. Tübingen: Narr, 29–43.
- Merisuo-Storm, Tuula & Soininen, Marjaana (2014). Students' first language skills after six years in bilingual education. *Mediterranean Journal of Social Sciences* 22(5), 72–81.
- Meyer, Hilbert (2005). *Was ist guter Unterricht?* Berlin: Cornelsen Scriptor.
- Meyer, Christiane (2003). *Bedeutung, Wahrnehmung und Bewertung des bilingualen Geographieunterrichts. Studien zum zweisprachigen Erdkundeunterricht (Englisch) in Rheinland-Pfalz*. URL: <http://ubt.opus.hbz-nrw.de/volltexte/2004/192/pdf/20021118.pdf>, Zugriff am 09.12.2014.
- Neumann, Marco; Milek, Anne; Maaz, Kai & Gresch, Cornelia (2010). Zum Einfluss der Klassenzusammensetzung auf den Übergang von der Grundschule in die weiterführenden Schulen. In Kai Maaz, Jürgen Baumert, Cornelia Gresch & Nele McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Bonn: BMBF, 229–251.
- Nold, Günter; Hartig, Johannes; Hinz, Silke & Rossa, Henning (2008). Klassen mit bilinguaem Sachfachunterricht: Englisch als Arbeitssprache. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz, 451–457.
- O'Connell, Ann A. & McCoach, Betsy D. (Hrsg.) (2007). *Multilevel modeling of educational data*. Charlotte, NC: Information Age Publishing.
- Otten, Edgar & Wildhage, Manfred (2007). Content and Language Integrated Learning. Eckpunkte einer „kleinen“ Didaktik des bilingualen Sachfachunterrichts. In Manfred Wildhage & Edgar Otten (Hrsg.), *Praxis des bilingualen Unterrichts*. Berlin: Cornelsen, 12–45.
- Pandel, Hans-Jürgen (2005). *Geschichtsunterricht nach Pisa: Kompetenzen, Bildungsstandards und Kerncurricula*. Schwalbach/ Ts.: Wochenschau Verlag.
- Pérez-Cañado, María (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism* 15(3), 315–341.

- Piesche, Nicole; Jonkmann, Kathrin & Keßler, Jörg-U. (2014). „Effekte des bilingualen Unterrichts auf das sachfachliche Lernen im naturwissenschaftlichen Unterricht - ein Feldexperiment“. Posterbeitrag auf der zweiten Tagung der Gesellschaft für empirische Bildungsforschung (GEBF) in Frankfurt am 03.03.2014.
- PISA-Konsortium (Hrsg.) (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Poropat, Arthur E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin* 135(2), 322–338.
- Raudenbush, Stephen W. & Bryk, Anthony S. (2002). *Hierarchical linear models*. Thousand Oaks: Sage.
- Rautenhaus, Heike (2010). Prolegomena zu einer Didaktik des bilingualen Sachfachunterrichts. Beispiel: Geschichte. In Gerhard Bach & Susanne Niemeier (Hrsg.), *Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven*. Frankfurt/ Main: Lang, 109–120.
- Ruiz de Zarobe, Yolanda (2010). Written production and CLIL. An empirical study. In Christiane Dalton-Puffer, Tarja Nikula & Ute Smit (Hrsg.), *Language use and language learning in CLIL classrooms*. Amsterdam: Benjamins, 191–209.
- Rumlich, Dominik (2014a). Prospective CLIL and non-CLIL students' interest in English (classes): A quasi-experimental study on German sixth-graders. In Ruth Breeze, Concepción Pasamar, Carmen Saíz & Cristina Sala (Hrsg.), *Integration of theory and practice in CLIL*. Amsterdam: Rodopi, 75–95.
- Rumlich, Dominik (2014b). „Selektion und Präparation als Ursache für die hohen Sprachkompetenzen bilingual unterrichteter Lerner? Kritische Evaluation einer Längsschnittstudie“. Unveröffentlichter Vortrag auf der zweiten Tagung der Gesellschaft für empirische Bildungsforschung (GEBF) in Frankfurt am 05.03.2014.
- Rumlich, Dominik (2012). Anwendungsbeitrag. (Sprach-)Tests in der Praxis: Die Studie „Development of North Rhine-Westphalian CLIL Students“. In Sabine Doff (Hrsg.), *Fremdsprachenunterricht empirisch erforschen: Grundlagen – Methoden – Anwendung*. Tübingen: Narr, 169–178.
- Saalbach, Henrik; Eckstein, Doris; Andri, Nicoletta; Hobi, Reto & Grabner, Roland H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction* 26, 36–44.
- Schwab, Götz (2013). Bili für alle? Ergebnisse und Perspektiven eines Forschungsprojekts zur Einführung bilingualer Module in einer Hauptschule. In Stephan Breidbach & Britta Viebrock (Hrsg.), *CLIL: Research, policy and practice*. Frankfurt/ Main: Lang, 297–314.
- Snow, Richard (1989). Aptitude-treatment interaction as framework for research on individual differences in learning. In Phillip Ackermann, Robert J. Sternberg & Robert Glaser (Hrsg.), *Learning and individual differences*. New York: W. H. Freeman, 13–59.

- Steinlen, Anja & Piske, Thorsten (2015). Zur Entwicklung der Schulleistungen von Kindern mit und ohne Migrationshintergrund in einer bilingualen Grundschule. Eine Pilotstudie. In Markus Kötter & Jutta Rymarczyk (Hrsg.), *Englisch auf der Primarstufe. Neue Forschungen – Weitere Entwicklungen*. Frankfurt/ Main: Lang, 123–149.
- Sylvén, Liss K. (2010). *Teaching in English or English teaching? On the effects of Content and Language Integrated Learning on Swedish learners' incidental vocabulary acquisition*. Göteborg: Acta Universitatis Gothoburgensis.
- Theis, Rolf (2010). Bilingualer Geschichtsunterricht. In Sabine Doff (Hrsg.), *Bilingualer Sachfachunterricht in der Sekundarstufe. Eine Einführung*. Tübingen: Narr, 44–57.
- Thürmann, Eike (2013). Spezifische Methoden für den bilingualen Unterricht/ CLIL. In Wolfgang Hallet & Frank G. Königs (Hrsg.), *Handbuch Bilingualer Unterricht. Content and Language Integrated Learning*. Seelze: Kallmeyer, 229–235.
- Viebrock, Britta (2007). *Bilingualer Erdkundeunterricht. Subjektive didaktische Theorien von Lehrerinnen und Lehrern*. Frankfurt/ Main: Lang.
- Zydatið, Wolfgang (2012). Linguistic thresholds in the CLIL classroom? The threshold hypothesis revisited. *International CLIL Research Journal* 4(1), 17–28. URL: <http://www.icrj.eu/14/article2.html/>, Zugriff am 09.12.2014.