

Bei dieser Arbeit handelt es sich um eine Wissenschaftliche Hausarbeit, die an der Universität Kassel angefertigt wurde. Die hier veröffentlichte Version kann von der als Prüfungsleistung eingereichten Version geringfügig abweichen. Weitere Wissenschaftliche Hausarbeiten finden Sie hier: <https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2011040837235>

Diese Arbeit wurde mit organisatorischer Unterstützung des Zentrums für Lehrerbildung der Universität Kassel veröffentlicht. Informationen zum ZLB finden Sie unter folgendem Link:

www.uni-kassel.de/zlb

Universität Kassel

Wissenschaftliche Hausarbeit zur Ersten Staatsprüfung für das Lehramt an Gymnasien im
Fach Biologie, eingereicht dem Amt für Lehrerbildung – Prüfungsstelle Kassel –

**Einfluss der kognitiven Leistungsfähigkeit von
Schülerinnen und Schülern auf die Bewertungskompetenz
im bilingualen Biologieunterricht**

Verfasser: Daniel Busch

Gutachter: Prof. Dr. Jürgen Mayer
FB 10 • Naturwissenschaften und Mathematik
AG Didaktik der Biologie

01.03. – 21.05.2012
Verlängerung: 21.06.2012

Inhaltsverzeichnis

Abbildungsverzeichnis.....	4
Tabellenverzeichnis	5
1 Einleitung.....	7
2 Theoretische Grundlagen.....	9
2.1 Organisatorische Formen des bilingualen Unterrichts.....	9
2.1.1 Definition wichtiger Begriffe und Modelle des bilingualen Unterricht.....	9
2.1.2 Stand und Entwicklung von Bilinguaem Unterricht – Organisation und aktuelle Zahlen...	13
2.1.3 Biologie als bilinguales Unterrichtsfach.....	16
2.2 Die Cognitive Load Theory	18
2.2.1 Beschreibung der <i>Cognitive Load Theory</i>	18
2.2.2 Ausgewählte Effekte der <i>Cognitive Load Theory</i> auf den Lernprozess	24
2.2.3 Anpassung der Theorie an den bilingualen Biologieunterricht	26
2.3 Der Kompetenzbereich Bewerten.....	28
2.3.1 Bewertungskompetenz im Zusammenhang mit den Bildungsstandards.....	28
2.3.2 Bewerten im bilingualen Biologieunterricht.....	33
3 Fragestellung.....	33
3.1 Formulierung der Fragestellung.....	33
3.2 Hypothesen	35
4 Methodisches Vorgehen	37
4.1 Die Stichprobe	37
4.2 Die Messinstrumente	41
4.3 Die Durchführung der Datenerhebung.....	46
4.4 Die Auswertung	47
5 Ergebnisse.....	48
5.1 Allgemeine Ergebnisse	48
5.1.1 Anmerkungen zum Kognitionstest und Ergebnisse	48
5.1.2 Ergebnisse der Testheftaufgaben	51
5.1.3 Ergebnisanalyse hinsichtlich der Teilkompetenzen nach dem Oldenburger Modell von REITSCHERT ET AL.	57
5.2 Ergebnisse zu den Unterfragestellungen.....	62
5.2.1 Unterscheidung nach Alter bzw. Jahrgangsstufe.....	63
5.2.2 Unterscheidung zwischen Biologie bilingual und Biologie auf Deutsch.....	74
5.2.3 Unterscheidung zwischen bilinguaem Zug und bilinguaem Modul	83
5.2.4 Zusammenfassende Ergebnisse bezüglich der Korrelation von kognitiver Leistungsfähigkeit und Ausprägung der Bewertungskompetenz.	87

5.3. Ergebnisse Fragebogen-Items der bili-Teilnehmer.....	89
6 Diskussion und Ausblick	92
6.1 Diskussion der Ergebnisse	92
6.1.1 Allgemeine Ergebnisse	92
6.1.2 Überprüfung der Hypothesen.....	95
6. 2 Bedeutung der Ergebnisse für den Biologieunterricht.....	104
6.2.1 Bedeutung der Ergebnisse für den bilingualen Biologieunterricht	104
6.2.2 Bedeutung der Ergebnisse für den Erwerb der Bewertungskompetenz im Biologieunterricht	106
6.3 Methodenkritik	107
7 Literaturverzeichnis	111
Erklärung	116
Anhang.....	117

Abbildungsverzeichnis

Abbildung 1 Die drei Kategorien der Cognitive Load (nach CLT, SWELLER 2005).....	21
Abbildung 2: Teilkompetenzen von Bewertungskompetenz (nach REITSCHERT ET AL. 2007, aus MITTELSTEN SCHEID & HÖBLE 2007).....	31
Abbildung 3: Verteilung Bildungsgang (in Prozent) der Teilnehmer der Studie (n=169; k.A.= keine Angabe).....	38
Abbildung 4: Absolute Altersverteilung der Studienteilnehmer (n=169).....	39
Abbildung 5: Geschlechterverteilung (in Prozent) der Teilnehmer der Studie (n=169).....	39
Abbildung 6: Verteilung der Unterrichtssprache (in Prozent) der Teilnehmer der Studie (n=169)	39
Abbildung 7: Durchschnittliche Notenpunkte der Studienteilnehmer nach Biologieunterricht; 15-Punkte-Skala (n=169)	40
Abbildung 8: Anzahl der Teilnehmer pro ermittelte Überlastungsschwelle (n=169).....	50
Abbildung 9: Anzahl der Antworten der Teilnehmer pro Niveaustufe für die Aufgabenteile 1A und 5D (n=169)	51
Abbildung 10: Anzahl der Antworten der Teilnehmer pro Niveaustufe für die Aufgabenteile 1D (n=169)	53
Abbildung 11: Ergebnisse der Aufgaben 2A, 2B, 3 und 4A (n=169).....	55
Abbildung 12: Ergebnisse der Aufgabe 5, außer Teilaufgabe D; Verteilung der Ergebnisse auf die Niveaustufen (n=169)	57
Abbildung 13: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei den 15-16-Jährigen (n=23)	68
Abbildung 14: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei den 17-18-Jährigen (n=136)	69
Abbildung 15: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei den 19-21-Jährigen (n=10)	69
Abbildung 16: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern des gymnasialen Bildungsganges G8 (n=60)	73
Abbildung 17: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern des gymnasialen Bildungsganges G9 (n=101)	73
Abbildung 18: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern mit bilinguaem Biologieunterricht (n=105).....	77
Abbildung 19: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern mit Biologie auf Deutsch (n=64).....	78
Abbildung 20: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern nach Unterrichtssprache und Alter.	82
Abbildung 21: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern in einem bilingualen Zug (n=51).....	86
Abbildung 22: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern in einem bilingualen Modul (n=53).....	87
Abbildung 23: Gesamtkorrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz (n=169)	88
Abbildung 24: Durchschnittliche Zustimmung je Item (n=104)	90
Abbildung 25: Zustimmung zu Items 3 und 4 geordnet nach Überforderungsschwelle (n=104).....	91

Tabellenverzeichnis

Tabelle 1: Anzahl der Schulen mit bilingualem Angebot in Englisch. HE= Hessen; D= Deutschland. Verändert nach WERNER 2005.	13
Tabelle 2: Definition der Niveaustufen zu den Antworten zu den Aufgabe 1A und 5D	51
Tabelle 3: Definition der Niveaustufen zu den Antworten zu der Aufgabe 1D.....	53
Tabelle 4: Ergebnisse der Teilaufgaben 5A, 5C und 5E.....	56
Tabelle 5: Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, nach Teilaufgaben; in %, höchster Wert markiert	59
Tabelle 6: Anteil der Stichprobe an den CREDIT-Wertungseinheiten nach Teilaufgaben; in %, höchster Wert markiert	59
Tabelle 7: Aussagekraft des Korrelationskoeffizienten r (nach Pearson).....	63
Tabelle 8: Durchschnittliche Kognitions- und Kompetenzwerte für die drei Altersgruppen	64
Tabelle 9: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1A, 1B, 1C, 1D; höchster Wert markiert. Aufteilung nach Altersgruppen.....	65
Tabelle 10: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1E, 4B, 5A, 5B; höchster Wert markiert. Aufteilung nach Altersgruppen	65
Tabelle 11: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5C, 5D, 5E, 5F; höchster Wert markiert. Aufteilung nach Altersgruppen	66
Tabelle 12: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungseinheiten von Bewertungskompetenz, Aufgabe 2, 3, 4A; höchster Wert markiert. Aufteilung nach Altersgruppen...	67
Tabelle 13: Durchschnittliche Kognitions- und Kompetenzwerte für die G8-Gruppe und die G9-Gruppe	70
Tabelle 14: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1 und 4B, höchster Wert markiert. Aufteilung nach Bildungsgang.....	71
Tabelle 15: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5, höchster Wert markiert. Aufteilung nach Bildungsgang	71
Tabelle 16: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungseinheiten von Bewertungskompetenz, Aufgabe 2, 3, 4A; höchster Wert markiert. Aufteilung nach Bildungsgang ...	72
Tabelle 17: Durchschnittliche Kognitions- und Kompetenzwerte für die bili-Gruppe und die mono-Gruppe	74
Tabelle 18: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1 und 4B höchster Wert markiert. Aufteilung nach Biologieunterricht.....	75
Tabelle 19: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5; höchster Wert markiert. Aufteilung nach Biologieunterricht	76
Tabelle 20: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungen der Bewertungskompetenz, Aufgabe 2, 3 und 4A; höchster Wert markiert. Aufteilung nach Biologieunterricht.....	76
Tabelle 21: Durchschnittliche Kognitions- und Kompetenzwerte für die Gruppen bG1, bG2, mG1 und mG2	78
Tabelle 22: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1B-1D; höchster Wert markiert. Aufteilung nach Biologieunterricht und Alter	79
Tabelle 23: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5A, 5C und 5F; höchster Wert markiert. Aufteilung nach Biologieunterricht und Alter	80
Tabelle 24: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungen Bewertungskompetenz, Aufgabe 2; höchster Wert markiert. Aufteilung nach Biologieunterricht und Alter	81
Tabelle 25: Korrelationswerte der Gruppen bG1, bG2, mG1 und mG2.....	83
Tabelle 26: Durchschnittliche Kognitions- und Kompetenzwerte für die Zug-Gruppe und die Modul-Gruppe	83

Tabelle 27: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5; höchster Wert markiert. Aufteilung nach bilingualem Zug (n=51) und bilingualem Modul (n=53).	84
Tabelle 28: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5; höchster Wert markiert. Aufteilung nach bilingualem Zug (n=51) und bilingualem Modul (n=53).	85
Tabelle 29: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungen der Bewertungskompetenz, Aufgabe 2, 3 und 4A; höchster Wert markiert. Aufteilung nach bilingualem Zug (n=51) und bilingualem Modul (n=53)	85
Tabelle 30: Korrelationswerte von kognitiver Leistungsfähigkeit und Bewertungskompetenz für die verschiedenen Stichproben	88
Tabelle 31: Zustimmung zu Items Motivation, Überforderung Sprache, Überforderung Inhalt und Bili-Wahl bereut, nach Kompetenzwertgruppen (1= Stimmt genau; 5= Stimmt gar nicht; n=104)	91
Tabelle 32: Korrelationswerte für die Teilkompetenzen „Urteilen/Argumentieren“ und „Ethisches Basiswissen“ für die gesamte Stichprobe, sowie die bili-Gruppe und die mono-Gruppe	98

1 Einleitung

Nach dem überraschend schlechten Abschneiden deutscher Schülerinnen und Schüler bei dem PISA-Test der OECD-Länder im Jahr 2000, treffenderweise auch als PISA-Schock bezeichnet, ging ein Umdenken durch die deutsche Bildungslandschaft. Die schulischen Rahmenbedingungen, die Lehrinhalte und die Curricula wurden überdacht und überarbeitet. Im Fach Biologie wurden durch die Kultusministerkonferenz verbindliche Bildungsstandards und Kompetenzen implementiert. Die vier Kompetenzbereiche Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewerten definieren mittlerweile den Biologieunterricht der Mittelstufe, während das Kerncurriculum für die Oberstufe Ähnliches für die gymnasiale Oberstufe darstellt.

Gleichzeitig wurde bilingualer Unterricht vermehrt in den Fokus gestellt. Die Zahl der Schulen, die Sachfachunterricht auf Englisch anbieten ist im letzten Jahrzehnt stark gestiegen und bilingualer Unterricht wurde damit zu einem der entscheidenden Faktoren für die Profilierung von Schulen in Deutschland. Von diesen beiden Maßnahmen versprechen sich Erziehungswissenschaftler, Didaktiker und Bildungspolitiker viel. Die Bildungsstandards sollen die großen Mängel des deutschen Bildungssystems ausmerzen, während bilingualer Unterricht die Schülerinnen und Schüler in Zeiten der Globalisierung auf internationale Arbeit vorbereiten und ihnen Kulturverständnis vermitteln soll. Dies sind sicherlich hehre Ziele, allerdings steht ein Nachweis der erhofften positiven Auswirkungen zurzeit noch aus. Besonders bilingualer Unterricht wird durchweg von Didaktikern, zumeist von Fremdsprachendidaktikern, gelobt. Beinahe fundamentalistisch-optimistisch werden die Vorteile von bilinguaem Unterricht herausgestellt, negative Stimmen hört und liest man hingegen selten.

Der bilinguale Unterricht im Fach Biologie orientiert sich inhaltlich am Curriculum des Sachfaches. Die englische Sprache dient, entgegen der Befürchtung vieler Biologielehrer, nur als Arbeitssprache. Folglich gelten die Bildungsstandards im Fach Biologie eben auch für den bilingualen Unterricht, schließlich ist die Minimalforderung, die für bilingualen Unterricht gilt, dass Schülerinnen und Schüler, die diesen belegen, mindestens genauso viel Fachliches lernen sollen wie ihre Mitschüler, die das Sachfach in der Muttersprache belegen. Befürworter von

bilingualem Biologieunterricht betonen regelmäßig, dass sich gerade dieses Fach für den englischsprachigen Unterricht eigne. Viele Fachbegriffe in den Naturwissenschaften basierten durch die Wissensglobalisierung sowieso schon auf englischsprachigen Begriffen und würden von diesen abgeleitet oder schlicht übernommen. Sicherlich sind das korrekte Aussagen und im Kompetenzbereich Fachwissen kommt den Schülerinnen und Schülern diese begriffliche Überschneidung zugute. Im Kompetenzbereich Bewerten ist das allerdings nicht unbedingt der Fall. Schülerinnen und Schüler müssen nach den Vorstellungen der KMK in der Lage sein, Urteile in ethisch-moralischen Kontexten und in solchen, die die nachhaltige Entwicklung betreffen, zu fällen. Dieser Typus von Anforderung ist einer, bei dem sich die deutschen Schülerinnen und Schüler bei PISA nicht gerade ausgezeichnet haben. Aufgaben und Unterrichtsbeispiele, die eine Entscheidung in einem ethischen Dilemma anstreben, bedeuten eine hohe kognitive Belastung für die Schülerinnen und Schüler. Kognitive Belastung, beziehungsweise Überlastung, ist auch ein zentraler Aspekt der *Cognitive Load Theory*, die mittlerweile die zentrale Theorie zum Lernen und kognitiven Vorgängen darstellt. Die Theorie besagt grob, dass die kognitive Leistungsfähigkeit jedes Menschen durch eine natürliche Begrenzung des Arbeitsgedächtnisses limitiert wird. Erhält ein Lerner einen zu hohen Input, das heißt, stellt eine Aufgabe eine zu hohe kognitive Anforderung dar, so wird das Arbeitsgedächtnis überlastet und ein Lernprozess wird nicht, oder nur im geringen Maße, initiiert.

Die *Cognitive Load Theory* lässt sich ihrerseits auf den bilingualen Unterricht übertragen: komplexe Unterrichtsinhalte, wie beispielsweise das ethische Bewerten im Biologieunterricht weisen eine hohe inhärente Komplexität auf. Diese Arbeit geht der Frage nach, ob bilingualer Biologieunterricht für Schülerinnen und Schüler eine kognitive Überforderung darstellt, wenn komplexe Unterrichtsinhalte behandelt werden. Sollte dies der Fall sein, hätte das weitreichende Konsequenzen für den bilingualen Biologieunterricht. Würden die Inhalte bereits eine hohe Belastung auf den Lerner ausüben, müsste der Lehrer auf weniger komplexe, einfachere Vermittlungsmethoden zurückgreifen. Dies wiederum wäre nur schwer mit der Forderung der Wissenschaftspropädeutik und dem Erlernen der wissenschaftlichen Arbeitsweise vereinbar. Letztendlich würde diese Überforderung bedeuten, dass das oben erwähnte Minimalziel nicht erreicht werden kann. Zur Klärung der

Fragestellung liegt dieser Arbeit eine empirische Befragung von Schülerinnen und Schülern der gymnasialen Oberstufe zugrunde, die bilingualen Biologieunterricht besuchen. Mithilfe eines kognitiven Leistungstest und einem Test mit biologischen Aufgaben zum Kompetenzbereich Bewerten wird quantitativ ein Einfluss der kognitiven Leistungsfähigkeit von Schülerinnen und Schülern auf die Bewertungskompetenz im bilingualen Biologieunterricht untersucht.

2 Theoretische Grundlagen

2.1 Organisatorische Formen des bilingualen Unterrichts

Obwohl bilingualer Unterricht (BU), das heißt der Unterricht in einem Sachfach wie eben Biologie auf Englisch oder einer anderen Fremdsprache, mittlerweile ein fester Bestandteil der deutschen Bildungslandschaft ist, steckt die zugehörige Didaktik immer noch in den Kinderschuhen. Es fehlt ein eigenständiger Lehrplan für das jeweilige bilingual unterrichtete Fach, sowie sinnvolle Synergien zwischen den Sachfächern und dem Lehrplan für das Unterrichtsfach Englisch. Besonders die Fachdidaktiker nehmen sich aus diesem Entwicklungsprozess heraus und überlassen den Fremdsprachendidaktikern das Feld. Dass es in der Folge zu Unmut über Gewichtigungen und Einfluss kommt, ist daher nicht weiter verwunderlich. Ein großes Problem ist, dass zwar viele Lehrer, Schulleiter und Eltern bilingualen Unterricht gutheißen und fordern, ihnen aber oftmals nicht klar ist, was dieses Konzept eigentlich bedeutet. Im Folgenden werden daher die wichtigsten Definitionen und Modelle des bilingualen Unterrichts diskutiert. Es gibt in Deutschland bilinguale Ansätze, die verschiedene Fremdsprachen als Arbeitssprache verwenden, aber Sachfächer auf Englisch bilden die deutliche Mehrheit, weshalb in Folgenden nur deutsch-englische Ansätze betrachtet werden. Mehr zur Organisation von bilingualem Unterricht wird weiter unten dargestellt.

2.1.1 Definition wichtiger Begriffe und Modelle des bilingualen Unterricht

Der Begriff des bilingualen Unterrichts an sich ist zuerst einmal missverständlich. In der Regel wird im Unterricht auf Englisch gesprochen, nur in Ausnahmefällen und im Anfangsunterricht wird häufiger auf die deutsche Sprache zurückgegriffen. Dennoch sollen Schülerinnen und Schüler, die ein Sachfach bilingual belegen, auch die deutschen Fachtermini lernen. Hier ist also sogenanntes *Code Switching* (vgl.

FEHLING 2008) unabdinglich. In den europäischen Nachbarländern hat sich hingegen der Begriff *Content and Language Integrated Learning* (CLIL) durchgesetzt und wird auch in Fachdiskussionen und -publikationen verwendet. Dieser Begriff passt deutlich besser, da er das Wesen des bilingualen Unterrichts erläutert: fachliche und sprachliche Inhalte werden im Unterricht integriert, wobei das fachliche Lernen (Content) die Hauptrolle spielt, sprachliches Lernen jedoch einen wesentlichen Anteil ausmacht.

Eine der Hauptkritikpunkte der Fachdidaktiker gegen bilingualen Unterricht befasst sich mit der Angst, dass der Fachunterricht zu einem erweiterten Sprachenunterricht verkommt, der zu Lasten der fachlichen Inhalte geht. Dieser Kritik muss deutlich widersprochen werden. Der Lehrplan des Sachfaches stellt den Orientierungsleitfaden für die Inhalte des bilingualen Unterrichts dar. Die englische Sprache wird als Arbeitssprache verwendet. Man könnte den bilingualen Unterricht in dieser Hinsicht mit einem Auslandsstudium vergleichen. Wenn ein Student in einem englischsprachigen Land für ein oder mehrere Semester Biologie studiert, laufen natürlich auch gewisse, nicht zu verachtende Sprachlernprozesse ab. Dennoch würde niemand dieses Auslandsstudium als verkapptes Englischstudium bezeichnen. Vielmehr würden die Möglichkeiten der Sprachimmersion und des natürlichen Spracherwerbs der Fachsprache, aber auch der Alltagskommunikation als positiver Nebeneffekt gelobt werden.

Im bilingualen Unterricht ist genau dies möglich. Die Schülerinnen und Schüler, die bilingualen Unterricht belegen, profitieren nicht nur fachlich, sondern auch sprachlich. Die Fremdsprachendidaktikerin SYLVIA FEHLING untersucht in einer empirischen Studie die Entwicklung der *Language Awareness*, die ein bedeutender Indikator für die Beherrschung einer Fremdsprache ist, und kommt zum Ergebnis, dass bilingual unterrichtete Schüler einerseits eine höhere totale *Language Awareness* aufwiesen, andererseits aber auch einen höheren Zuwachs über zwei Jahre aufwiesen (vgl. FEHLING 2009). Die Unterschiede erwiesen sich als signifikant. Auch die DESI-Studie (vgl. DESI-KONSORTIUM 2008) belegt einen sprachlichen Vorteil der bilingual unterrichteten Schülerinnen und Schülern bei der sprachlichen Kompetenz für Englisch. Insgesamt wurde ein Vorsprung von bis zu zwei Schuljahren in den verschiedenen sprachlichen Teilkompetenzen nachgewiesen.

Auf den ersten Blick scheinen diese Befunde den Kritikern recht zu geben, besonders weil es kaum empirische Befunde dafür gibt, dass Schülerinnen und Schüler, die bilingualen Sachfachunterricht besuchen, auch bei fachlichen Inhalten ihren muttersprachlich unterrichteten Mitschülern überlegen sind. Auf den zweiten Blick muss allerdings auffallen, dass die sprachliche Kompetenz, die im BU erlangt wird, nicht mit der zu vergleichen ist, die aus dem Fremdsprachenunterricht resultiert. Letzterer ist vorrangig auf die Fähigkeit, sinnvoll mit Muttersprachlern kommunizieren zu können, ausgerichtet und wird auch dahingehend bewertet. Der Lehrplan Englisch für den gymnasialen Bildungsgang definiert die Ziele des Englischunterrichts als „kommunikative Kompetenz“ und weist darauf hin, dass Englisch „in vielen Ländern Muttersprache, in zahlreichen Ländern zumindest eine der Amtssprachen, in den meisten Ländern die vorwiegend unterrichtete erste Fremdsprache [ist]“ (HESS. KULTUSMINISTERIUM 2010, S. 2). Es geht dabei um alltägliche Kommunikation zwischen Individuen. Günter Nold et al. kritisieren in den Ergebnissen der DESI-Studie, dass sich „durch die Beschränkung des Gebrauchs der Fremdsprache auf den Englischunterricht (im Falle von Englisch) eine umfassende sprachlich-kommunikative und pragmatisch-interkulturelle Kompetenz nur in Ausnahmefällen erreichen [lässt]“ (NOLD et al. 2008, S. 452). Im Sachfach dagegen, und dies ist auch im muttersprachlichen Unterricht der Fall, steht die fachliche Kommunikation im Kontext der Inhalte des betreffenden Fachbereichs im Vordergrund. Die Formulierung des Kompetenzbereiches Kommunikation durch die Kultusministerkonferenz, die sowohl in den Bildungsstandards für den Mittleren Schulabschluss, als auch im Kerncurriculum Oberstufe verankert sind (vgl. HARMS ET AL. 2004; KMK 2005), bestätigen dieses. JIM CUMMINS unterscheidet diese beiden Sprachfertigkeiten in seinem Konzept von BICS und CALP (vgl. CUMMINS 1980). BICS (*Basic Interpersonal Communicative Skills*) sind jene Fähigkeiten, die im Fremdsprachenunterricht erlernt werden, während CALP (*Cognitive Academic Language Proficiency*) als jene Sprachfertigkeiten definiert sind, die „über die reine face-to-face Kommunikation hinausgehen und komplexer, präziser und expliziter sind“ (FENKERT 2009, S. 18).

Ein weiterer Vorteil des bilingualen gegenüber dem muttersprachlichen Unterricht ist die Integration von interkulturellem Lernen. Durch die Verwendung von Unterrichtsmaterialien aus englischsprachigen Ländern ist immer auch ein Einblick

in eine fremde Kultur möglich. Gerade bei Themen, die den Kompetenzbereich Bewertung betreffen, sei es Stammzellforschung, Organspende, nachhaltige Entwicklung, Umweltschutz, Abtreibung, PID, etc., bietet der bilinguale Unterricht Möglichkeiten zum Perspektivwechsel an, die muttersprachlicher Unterricht nur schwer herstellen kann. Der Aspekt des interkulturellen Lernens ist für die Fragestellung dieser Arbeit allerdings nur von Randbedeutung, wird aus Gründen der Vollständigkeit an dieser Stelle erwähnt.

Neben dem sprachlichen Zugewinn wird in der Fachliteratur oft auch von einem Zuwachs an Motivation gesprochen. BOHN erklärt dieses Phänomen damit, dass „die Verwendung der Fremdsprache [...] von den Schülern als besonderer Reiz wahrgenommen [wird] und sie fordert ein konzentriertes Arbeiten und Lernen“ (2008, S. 30). HEMMELGARN UND EWIG argumentieren hingegen, dass bilinguale Kurse oftmals aus besonders leistungsstarken und motivierten Schülerinnen und Schülern zusammengesetzt ist und dass keine oder keine signifikanten Unterschiede bei der Motivation zu verzeichnen sind. Sie warnen im Gegenteil sogar vor einer Überstigmatisierung und einer Übererwartung an bilinguale Klassen, die einen gegenteiligen Effekt haben kann (vgl. HEMMELGARN & EWIG 2003). Es bleibt also die Frage, ob bilingual unterrichtete Lerngruppen durch den BU leistungsstärker und motivierter sind, oder ob sie BU belegen, weil sie von vornherein diese beiden Eigenschaften aufweisen. Noch ist bilingualer Unterricht als Phänomen aber zu jung, als dass über diese Frage endgültig entschieden werden könnte.

Die angesprochene Unklarheit darüber, ob Schülerinnen und Schüler, die BU belegen, auch fachlich vergleichbar gute Lernerfolge vorweisen können, stellt allerdings immer noch das größte Problem des bilingualen Unterrichts dar. Aufgrund eines eklatanten Mangels an empirischen Belegen kann eine fachliche Benachteiligung beim BU noch nicht ausgeschlossen werden. „Vor diesem Hintergrund erscheint es deshalb durchaus bedenklich [...], dass im Rahmen einer Befragung von 80 Bili-Lehrern 47% nicht ausschließen wollen, dass es im Verlauf des fremdsprachlichen Unterrichts zu fachlichen Defiziten kommen kann“ (BOHN 2008:35). Zwar bezieht sich diese Aussage vorrangig auf die Breite der Darstellung und eben nicht auf die Tiefe, aber eine gewisse Skepsis bleibt. Das Minimalziel von BU muss daher lauten, dass Schülerinnen und Schüler, die diesen belegen, mindestens genauso viele fachliche Kompetenzen erlernen, wie ihre einsprachigen

Mitschüler. Dieses Ziel kann nur durch weitere Unterrichtsforschung und durch eine kontinuierliche, begleitende Aus- und Weiterbildung zur Qualifikation von Lehrerinnen und Lehrern zum BU erreicht werden. An diesem Punkt soll auch diese Arbeit ansetzen, da der Vergleich zwischen den fachlichen Leistungen von bilingual und monolingual unterrichteten Schülerinnen und Schülern Gegenstand der Fragestellung ist.

2.1.2 Stand und Entwicklung von Bilinguaem Unterricht – Organisation und aktuelle Zahlen

Bilingualer Unterricht erlebt seit einigen Jahren ein rapides Wachstum, wenngleich das Konzept schon fast 50 Jahre alt ist. 1963 unterzeichneten Deutschland und Frankreich zur Stärkung der freundschaftlichen Nachbarschaft den Elysée-Vertrag. Ein Resultat dieses Vertrages war die Einführung bilingualer Programme in beiden Ländern, in Deutschland bevorzugt im Grenzgebiet zum Nachbarn. Lange Zeit dominierten hernach französischsprachige Programme die bilinguale Bildungslandschaft. Der Vertrag von Maastricht 1992 hatte gut 30 Jahre später einen ähnlichen Effekt auf die Einrichtung deutsch-englischer Programme. Der heutige Boom geht also auf diesen Staatenvertrag zurück. Englisch hatte sich mittlerweile zur *lingua franca* und zur Wissenschaftssprache entwickelt, was der Entwicklung deutsch-englischer Angebote weiteren Auftrieb verschaffte.

Da Schulen die Einrichtung bilingualer Programme nicht an die Kultusministerien melden müssen, besonders dann nicht, wenn es sich nur um bilinguale Module handelt, ist es schwer, eine Aussage über die Verbreitung bilingualen Unterrichts in Deutschland zu machen. Man kann allerdings davon ausgehen, dass mittlerweile 400-500 Schulen im Bundesgebiet bilinguale Programme anbieten, Tendenz steigend. BETTINA WERNER zeichnet an Zahlen von 2005 nach, dass die Mehrheit der Programme in westdeutschen Schulen angesiedelt ist und dass besonders Gymnasien BU anbieten. Tabelle 1 stellt deutlich die steigende Anzahl der Schulen dar, die BU anbieten, allerdings unterscheidet WERNER nicht zwischen den verschiedenen Fächern, die jeweils angeboten werden.

Tabelle 1: Anzahl der Schulen mit bilinguaem Angebot in Englisch. HE= Hessen; D= Deutschland. Verändert nach WERNER 2005.

	Gymnasien		Gesamtschulen		Realschulen		Grundschulen	
	1993	2005	1993	2005	1993	2005	1993	2005

HE	1	21	1	25	-	11	-	-
D	81	349	8	75	34	70	-	5

Eine Veröffentlichung des hessischen Kultusministeriums zum bilingualen Angebot in Hessen im Jahr 2012 liefert exemplarisch einen Überblick über die aktuellen Entwicklungen. Im Moment bieten 42 Gymnasien, 30 Gesamtschulen und 13 Realschulen BU in Hessen an (vgl. HESS. KULTUSMINISTERIUM 2012). An den Zahlen wird deutlich, dass besonders die Angebote an Gymnasien sich stark vergrößert haben. Allerdings gibt auch diese Quelle nicht an, welche Fächer bilingual unterrichtet werden.

Des Weiteren unterscheiden die Autoren auch nicht zwischen bilingualen Schulen, bilingualen Zügen und bilingualen Modulen. In bilingualen Schulen, werden sämtliche Sachfächer zweisprachig unterrichtet. Diese Schulen spielen für die Fragestellung dieser Arbeit allerdings keine Rolle und können im Folgenden vernachlässigt werden. Die Unterscheidung zwischen Zügen und Modulen ist allerdings entscheidend, weil sie sich in einer Unterfragestellung der Arbeit wiederfindet. Schulen mit bilingualen Zügen bieten ihren Schülerinnen und Schülern schon früh an, bilingualen Sachfachunterricht zu belegen. Da hauptsächlich weiterführende Schulen BU im Schulprofil haben, ist Jahrgang 5 oder 6 der übliche Einstiegszeitpunkt für die Einwahl in einen bilingualen Zug. Zu Beginn wird ein Vorbereitungsschuljahr erteilt, bei dem der Englischunterricht um eine oder zwei Stunden aufgestockt wird. In der Regel sind Geschichte oder Geographie das erste Fach in Jahrgang 7, das auf Englisch unterrichtet wird. Sprachliche und inhaltliche Anfangsschwierigkeiten behindern den Lernprozess deutlich und die bilingualen Kurse benötigen mehr Zeit für die Behandlung der Stoffinhalte als ihre muttersprachlichen Mitschüler in den regulären Kursen. Um diesem Nachteil zu begegnen erteilen die Schulen oftmals eine zusätzliche Stunde Englischunterricht und eine weitere zusätzliche Stunde des betroffenen Sachfachs. Schülerinnen und Schüler, die einen bilingualen Zug besuchen, belegen das oder die Sachfächer, die sie auf Englisch besuchen, bis zum Schulabschluss als BU. Damit ist dieses Modell „durch ein Kontinuum gekennzeichnet, d.h. der bilinguale Unterricht beginnt meist mit einem bilingualen Vorbereitungskurs in Klasse 5 und 6 und endet Ende der Sekundarstufe I (Realschulbereich) bzw. der Sekundarstufe II (Gymnasium)“ (FEHLING 2008, S. 22). Zumindest in der Mittelstufe werden die bilingualen

Sachfächer durchgehend mit einer zusätzlichen Stunde unterrichtet, während der Englischunterricht nach dem ersten Jahr wieder auf die normale Stundenzahl zurückgefahren wird.

Wenn Schulen keine bilingualen Züge anbieten können oder wollen, können engagierte Lehrkräfte auf bilinguale Module zurückgreifen. Hierbei wird eine Unterrichtsreihe, ein Thema oder ein Halbjahr bilingual unterrichtet. Meistens gibt es für diesen Unterricht keine Regelungen der jeweiligen Schulverwaltung, weshalb es auch keine Anpassung der Stundentafel für das Sachfach und den Englischunterricht gibt. Dennoch sind diese Module eine populäre Methode, um Schülerinnen und Schüler mit bilinguaem Unterricht vertraut zu machen. Auch diese stark reduzierte Art von BU zeigt gute Ergebnisse und die Teilnehmer sind ihren Mitschülern oftmals voraus, wenn auch weit weniger deutlich, als dies bei bilingualen Zügen der Fall ist. In einer der Schulen wählen die Schülerinnen und Schüler vor dem Eintritt in die Oberstufe, ob sie Biologie bilingual belegen wollen. Zum Ende des Schuljahres steht dann wiederum eine Wahl an, ob die Schüler weiterhin den bilingualen Kurs besuchen wollen oder wieder Biologie auf Deutsch belegen. Somit stellt sich der BU Biologie an dieser Schule eher modular dar. In einer weiteren Schule, die eine reine Oberstufenschule ist, werden bilinguale Kurse erst in der Qualifikationsphase angeboten. In der Einführungsphase, in der Schülerinnen und Schüler, die aus zahlreichen Schulen an die diese Schule kommen, wird aus Gründen der Orientierung darauf verzichtet. Seit kurzem kommen pro Jahrgang zwei bilinguale Kurse zustande, allerdings können auch Schülerinnen und Schüler, die Biologie auf Deutsch gewählt haben, in bilinguale Kurse gelost werden, um überfüllte deutschsprachige Kurse zu verhindern. Auch in dieser Schule schwanken die Zahlen der Teilnehmer an BU in Biologie zwischen Q1/2 und Q3/4, weil zahlreiche Schüler doch wieder Biologie auf Deutsch belegen wollen oder Biologie ganz abwählen. In der dritten Schule sieht es ähnlich aus: in der Mittelstufe wird kein BU erteilt, ab der Oberstufe können die Schülerinnen und Schüler dann jedoch aus einem breiten Angebot von Sachfächern auswählen, sind aber nicht dazu verpflichtet. Die vierte Schule ist eine Europaschule, die in ihrem Konzept einen besonderen Fokus auf Handlungs- und Projektorientierung legt und das Ziel hat, die Schülerinnen und Schüler auf ein sinnstiftendes Leben in Europa vorzubereiten. Die Besonderheit dieser Schule im Kontext des bilingualen Unterrichts stellt die Organisationsform

bilingualer Züge dar. Die Schule hat 2003 einen bilingualen Zweig eingerichtet, in dem die Schüler in den Fächern Wirtschaftslehre, Politik, Geschichte, Mathematik, Sport und einer Naturwissenschaft bis zum Abitur in englischer Sprache unterrichtet werden.

2.1.3 Biologie als bilinguales Unterrichtsfach

Zwar ist Biologie das am häufigsten bilingual unterrichtete, naturwissenschaftliche Fach, allerdings gibt es im Vergleich zu den gesellschaftswissenschaftlichen Fächern Geographie, Geschichte und Politik auffallend wenig Angebote. HEMMELGARN und EWIG geben an, dass 1997 nur 3,8% der bilingualen Kurse auf das Fach Biologie entfielen (vgl. 2008). Das ist verwunderlich, eignet sich nach Meinung von Richter und Zimmermann (2003) doch gerade Biologie hervorragend für den BU: Englisch ist die vorherrschende Wissenschaftssprache und ohne Rückgriffe auf diese kann nicht mehr wissenschaftspropädeutisch gearbeitet werden. Das zeigt sich gerade daran, dass aktuelle Publikationen des Feldes durchgängig auf Englisch veröffentlicht werden. Des Weiteren werden neue Fachtermini zum großen Teil direkt aus dem anglophonen Raum übernommen. Ein Blick in den Lehrplan Biologie offenbart, dass ein Großteil der Inhalte erst in den letzten 40-50 Jahren entwickelt oder entdeckt worden sind, was den hohen Anteil anglophoner Fachbegriffe erklärt. Ältere Begriffe haben hingegen sowohl im Deutschen, als auch im Englischen meist lateinischen Ursprung (vgl. BOHN 2008). Hier besteht also eine große Überschneidung. „Es kommt hinzu, dass der Sprachduktus in der *scientific community* (sic) des angloamerikanischen Raumes sehr kommunikationsorientiert und auf Verständlichkeit ausgerichtet, also sehr stark an der Alltagssprache orientiert ist“ (Richter & Zimmermann 2003, S. 116). Einige Didaktiker sehen gerade darin einen besonderen Vorteil des bilingualen Biologieunterrichts. KRECHEL stellt fest, dass sich von allen Naturwissenschaften „in erster Linie Biologie an[bietet], weil auch hier im Anfangsunterricht der fachsprachliche Anteil durch allgemeinsprachliche Strukturen ersetzt werden kann“ (2008:4).

Ein weiteres schlagkräftiges Argument für bilingualen Biologieunterricht ist der Umstand, dass die Naturwissenschaften eine deutlich geringere nationale oder kulturelle Prägung aufweisen als die Gesellschaftswissenschaften. „Naturwissenschaftler müssen sich keine Gedanken darüber machen, ob sie ein

Thema [...] aus deutscher, amerikanischer oder italienischer Perspektive betrachten wollen und ob den deutschen Schülern bei der Verwendung von Unterrichtsmaterialien aus dem angelsächsischen Sprachraum eventuell spezifisch deutsche Betrachtungsweisen entgehen“ (BOHN 2008, S. 29). Dieses Argument kann im Umkehrschluss sogar noch stärker für BU in Biologie argumentieren, da es eben auch für Biologie in der Muttersprache gilt. Da aber bilingualer Biologieunterricht durch die Nutzung authentischer englischsprachiger Materialien auf Produkte einer anderen Kultur zurückgreift, ergibt sich eine seltene Chance auf interkulturelles Lernen im Biologieunterricht. Während nämlich die biologischen Grundlagen, wie beispielsweise die Durchführung eines genetischen Fingerabdrucks, wertfrei sind, trifft dies auf ihre Anwendung, z.B. im Falle der Präimplantationsdiagnostik, nicht zu. In der Diskussion um die Anwendung biologischer Methoden gibt es kulturspezifische Unterschiede und „deshalb bieten Themen wie der Umgang mit den modernen Reproduktionstechnologien in den USA oder Deutschland auch Ansätze zum Erwerb von Interkultureller Kompetenz“ (ibid.), die im muttersprachlichen Biologieunterricht so nicht erworben werden könnte.

Interessanterweise wird an den Schulen, die für diese Untersuchung herangezogen wurden, das Fach Biologie nur in der Oberstufe erteilt. Bei Schulen, die langfristig angelegte, bilinguale Züge anbieten, ist es somit oftmals erst das zweite oder dritte Fach, das bilingual unterrichtet wird. Dadurch haben diese Schülerinnen und Schüler natürlich einen klaren Vorteil gegenüber solchen, die nur modular BU belegen. Das Fach Geographie, das oftmals als bilinguales Einstiegsfach erteilt wird, weist große Überschneidungen mit dem Fachbereich Ökologie auf. Schülerinnen und Schüler, die im Rahmen eines bilingualen Zuges im 11. Jahrgang mit Biologie anfangen, können daher auf Vorwissen zurückgreifen und sehen sich kleineren sprachlichen Barrieren ausgesetzt als Schülerinnen und Schüler in modularem BU. Aber auch mit dem Lehrplan Englisch selbst gibt es zahlreiche Überschneidungen. HEMMELGARN UND EWIG (2003) haben einen Vergleich der beiden Lehrpläne aus NRW angestellt, der sich auf die hessischen Lehrpläne übertragen lässt. Die beiden Autoren vergleichen jeweils Aspekte der Lehrpläne und stellen Gemeinsamkeiten heraus. Für die hessischen Lehrpläne Biologie und Englisch (vgl. HESS. KULTUSMINISTERIUM 2010a; 2010b) könnte ein solcher Vergleich auch durchgeführt werden, würde allerdings den Umfang dieser Arbeit übersteigen. Exemplarisch sollen allerdings einige

Gemeinsamkeiten dargestellt werden. Parallelen bestehen in den Arbeits- und Unterrichtsformen; Partner- und Gruppenarbeit spielt in beiden Lehrplänen eine wichtige Rollen, aber auch das wissenschaftliche Arbeiten und Denken wird mit den Aspekten Hypothesenbildung- und prüfung in beiden Lehrplänen erwähnt. Auch die fachmethodische Verfahren der Darstellungsformen von Wissen auf visuelle, akustische, ganzheitliche und analytische Verfahren findet sich im Lehrplan Englisch. Inhaltlich findet sich beispielsweise der Respekt vor Lebewesen (Lehrplan Biologie) im interkulturellen Lernen (Lehrplan Englisch) wieder. Diese und die zahlreichen anderen Parallelen liefern ein weiteres Argument für bilingualen Unterricht.

2.2 Die Cognitive Load Theory

2.2.1 Beschreibung der *Cognitive Load Theory*

Die kognitions- und lernpsychologische Theorie, die dieser Arbeit zugrunde liegt, ist die *Cognitive Load Theory* (CLT) von JOHN SWELLER. Diese Theorie, die auf den Erkenntnissen von JOHN MILLER (vgl. 1956) beruht und aufbaut, besagt, dass alle kognitiven Vorgänge durch ein streng begrenztes Arbeitsgedächtnis limitiert sind. Je nach Aufgabentyp und Anforderung einer Aufgabe, die das Gehirn lösen muss, wird dieses Arbeitsgedächtnis belastet oder überlastet, was wiederum darin resultiert, dass kein Lernprozess abläuft. Die Theorie geht davon aus, dass Lernprozesse dann am Besten stattfinden, wenn die Rahmenbedingungen der Lernumgebung mit der kognitiven Architektur des Menschen zur Deckung gebracht werden. Ist dies nicht der Fall resultiert eben eine Über- oder manchmal auch eine Unterforderung. „Performance degrades at the cognitive load extremes of either excessively low load (underload) or excessively high load (overload) [...] Under conditions of both underload and overload, learners may cease to learn” (PAAS ET AL. 2004, S. 1). Dieser relative schmale Bereich zwischen Unter- und Überforderung macht eine Beachtung der CLT, sowie der Limitierung des Arbeitsgedächtnisses erforderlich, um langfristig erfolgreiche Lernumgebungen zu schaffen.

Das Modell von BADDELEY (1992) ist das bekannteste und einflussreichste Modell zum Arbeitsgedächtnis. Der Autor stellt darin fest, dass

the concept of a working memory system that temporarily stores information as part of the performance of complex cognitive tasks is proving to be productive. Studies that have utilized the individual difference approach have linked working memory to performance on a range of important tasks, including language comprehension and reasoning. [...] the concept forms a useful conceptual tool in understanding a range of neuropsychological deficits, which in turn have thrown light on normal cognitive functioning (BADDELEY 1992, S. 560).

Das Vorhandensein eines limitierten Arbeitsgedächtnisses wird heute von allen relevanten Forschern anerkannt und hat das Konzept des Kurzzeitgedächtnisses als kognitives Arbeitsinstrument verdrängt.

Laut MILLER wird die Limitierung des Arbeitsgedächtnisses mit dem abstrakten Begriff *Chunk* angegeben. Dieser Begriff wird aber von ihm nicht ausreichend definiert, sodass ein *Chunk* verschieden große Informationsmengen beinhalten kann. „The span of immediate memory seems to be almost independent of the number of bits per chunk, at least over the range that has been examined to date“ (MILLER 1956, S. 93). Folgt man dieser Aussage weiter, bedeutet sie, dass ein Lernprozess nichts anderes darstellt, als eine Recodierung von *Bits* in *Chunks* und eine Integration von *Chunks* in übergeordnete *Chunks*. MILLER verdeutlicht dieses am Beispiel eines Lernalters, der sich erstmalig mit dem Morsealphabet beschäftigt. Zuerst hört er nur eine scheinbar willkürliche Abfolge von kurzen und langen Tönen. Dann ist er in der Lage, diese willkürliche Abfolge als Buchstaben und später als Worte zu deuten. Er reduziert also fortlaufend die Anzahl an *Chunks*, indem er neue *Chunks* formt und neue Informationen in sein Vorwissen integriert. Kurz gesagt, „the operator learns to increase the bits per chunk“ (ibid.). Die Anzahl von *Chunks*, die vom Arbeitsgedächtnis zu einem Zeitpunkt verarbeitet werden kann, ist hingegen limitiert. Diese Zahl pendelt in vielen Fällen um die Zahl sieben, die von MILLER dann folglich auch „the magical number seven, plus or minus two“ genannt wird, weil das Limit unabhängig vom verwendeten Testformat, sowie von dem angesprochenen Sinn-Kanal (audiovisuell, visuell, akustisch, etc.) sich immer im Bereich von sieben Items befindet. Schon MILLER hat diese Information auf das Lernen bezogen, indem er vermutet, dass „as we increase the amount of input information, the observer will begin to make more and more errors; we can test the limits of accuracy of his absolute judgements“ (MILLER 1956, S. 82). Genau auf dieses Phänomen wird im kognitiven Leistungstest des Testheftes für diese Arbeit

zurückgegriffen (vgl. Kap. 4.2). Für die Beantwortung der biologierelevanten bilingualen Aufgaben bedeutet dies, dass diejenigen Aufgaben, die die kognitive Leistungsfähigkeit des individuellen Teilnehmers übersteigen, fehlerhaft oder im schlimmsten Fall gar nicht beantwortet werden können.

Auch die *Cognitive Load Theory*, die in den 1980er Jahren von JOHN SWELLER formuliert und in den folgenden drei Jahrzehnten durch zahlreiche Wissenschaftler weltweit (vgl. PAAS ET AL. 2003; PAAS ET AL. 2004; SWELLER 2005; KIRSCHNER ET AL. 2006) weiterentwickelt wurde, bezieht die Theorie des stark limitierten Arbeitsgedächtnisses mit ein. Sweller und Kollegen gehen davon aus, dass das Langzeitgedächtnis einen entscheidenden Einfluss auf alle kognitiven Vorgänge hat und dass Lernprozesse als Veränderungen der Informationen definiert werden können „as an alteration in long-term memory. If nothing has altered in long-term memory nothing has been learned“ (SWELLER 2005, S. 20). Die Konsequenzen dieser Aussage für die schulische Bildung sind eindeutig: Das Ziel des Unterrichts muss es sein, die im Langzeitgedächtnis gespeicherten Informationen zu (re-)aktivieren und die neuen Informationen, mit denen die Lernenden konfrontiert werden zu integrieren, zu assimilieren oder anzupassen. Das Langzeitgedächtnis weist hierbei, anders als das Arbeitsgedächtnis oder das Kurzzeitgedächtnis, kein Kapazitätslimit auf, zumindest konnte ein solches bislang nicht nachgewiesen werden.

Informationen, die im Langzeitgedächtnis gespeichert werden, werden als Schemata bezeichnet. SWELLER verdeutlicht, dass „Schemas are cognitive constructs that allow multiple elements of information to be categorized as a single element“ (SWELLER 2005, S. 21). Diese Beschreibung ist nahezu Deckungsgleich mit MILLERS Theorie der *Chunks*, nur dass SWELLER hier den Begriff *element* verwendet. Der Rückgriff auch vorhandene Schemata ist letztendlich auch die Lösung des Problems der Komplexität unserer Welt. Wäre ein Lernprozess, also eine Verpackung von unendlich vielen Informationen in *Chunks*, nicht möglich, wären wir konstant überfordert und wahrscheinlich kaum lebensfähig.

Die *Cognitive Load Theory* bezieht sich ursprünglich aber auf die Verarbeitung von neuartigen Informationen zu denen noch keine Schemata vorliegen. In solchen Situation greift der Mensch bei der Konfrontation mit einem Problem auf eine einfache *Trial-Error-Methode* zurück, um zu einer Lösung zu kommen. Diese

Methode ist zwar sehr uneffektiv und zeitaufwendig, stellt aber die einzige Möglichkeit zur Lösung des Problems dar. Wird das Problem letztendlich allerdings mit dieser Methode gelöst, so wird ein Schema angelegt und die oben beschriebene Veränderung des Langzeitgedächtnisses bewirkt, sodass ein Lernprozess stattgefunden hat. Dadurch bedingen sich das Langzeitgedächtnis und das Arbeitsgedächtnis auf verschiedene Art und Weise. Während das Arbeitsgedächtnis also eine Umstrukturierung der Informationen im Langzeitgedächtnis initiieren kann, verringert das Langzeitgedächtnis die Belastung des Arbeitsgedächtnisses, weil diese Informationen ohne Zusatzbelastung in das Arbeitsgedächtnis integriert werden kann. „Information that has already been organised into schemas in long-term memory can also be fed into working-memory. Neither the duration nor capacity limitations attached to novel information received from the sensory memory applies to information from long-term memory” (SWELLER 2005, S. 24). Bezieht man diesen Umstand mit ein, bedeutet es, dass das Arbeitsgedächtnis und seine Limitierungen irrelevant werden, wenn Schemata vorhanden und automatisiert sind (*Schema aquisition* und *Schema automatation*). Somit ergibt sich eine Hierarchie des Rückgriffs auf verschiedene Hilfssysteme, um Probleme zu lösen. Zuerst erfolgt der Rückgriff auf das eigene Langzeitgedächtnis. Weniger effektiv ist der Rückgriff auf Instruktionen einer dritten Person, die als Rückgriff auf ein fremdes Langzeitgedächtnis angesehen werden kann. Erst wenn beide Systeme keine zufriedenstellende Lösung generieren können, greift das Gehirn auf zufälliges Generieren von Lösungen mit anschließendem Effektivitätstest zurück.

Gehen wir aber jetzt wieder von neuartigen Informationen aus. Die daraus resultierende *Cognitive Load* setzt sich aus drei Kategorien von *Cognitive Load* zusammen (s. Abb.1).

$$\boxed{\text{Cognitive Load}} = \boxed{\text{Intrinsic Load}} + \boxed{\text{Extraneous Load}} + \boxed{\text{Germane Load}}$$

Abbildung 1 Die drei Kategorien der Cognitive Load (nach CLT, SWELLER 2005).

Die erste Kategorie wird *Intrinsic Cognitive Load* genannt und entzieht sich der instruktiven Kontrolle „due to [...]natural complexity of the information that must be processed. It is determined by levels of element interactivity“ (SWELLER 2005, S. 27). In verschiedenen Materialien gibt es eine verschieden hohe Interaktivität der

Elemente. Je höher diese ist, desto komplexer ist das Material und desto höher ist die *Intrinsic Load*. Die einzige Möglichkeit, die *Intrinsic Load* zu reduzieren ist der Austausch des Materials. Hierbei muss allerdings immer auch in Betracht gezogen werden, ob ein Austauschen oder Weglassen eines Materials ein Verstehen des Gesamtzusammenhanges gefährdet.

Die zweite Kategorie ist die *Extraneous Cognitive Load*, die auch *Ineffective Cognitive Load* genannt wird, weil sie für das Verstehen nicht notwendig ist und der Generierung und Automatisierung von Schemata im Wege steht (vgl. PAAS ET AL 2003). Diese Kategorie ist der Ansatzpunkt für Instruktion. Lehrende können an dieser Stellschraube drehen und den Lernprozess unterstützen, indem die *Extraneous Cognitive Load* möglichst klein gehalten wird. Die *Cognitive Load* Forschung hat ergeben, dass problemlösungsorientiertes Lernen eine sehr hohe *Extraneous Cognitive Load* bedeutet. „Any instructional procedure that requires learners to engage in either a search for a problem solution or a search for referents in an explanation [...] is likely to impose a heavy extraneous cognitive load because working memory resources must be used for activities that are irrelevant to schema acquisition and automation” (PAAS ET AL. 2003, S. 2). Diese Aussage ist wiederum nur schwer mit der wissenschaftspropädeutischen Forderung des modernen Biologieunterrichts vereinbar, der auch der bilinguale Unterricht unterworfen ist.

Die dritte Kategorie wird analog zur *Extraneous Cognitive Load* auch *Effective* oder *Germane Cognitive Load* genannt. „Whereas extraneous cognitive load interferes with learning, germane cognitive load enhances learning“ (ibid.). Diese Kategorie beschreibt sozusagen die gutartige kognitive Belastung, denn sie wird „caused by effortful learning resulting in schema construction and automation“ (SWELLER 2005, S. 27). Ein gutes Beispiel für *Germane Cognitive Load* sind zahlreiche Beispiele zu einem Phänomen. Zwar erhöhen die Beispiele die gesamte *Cognitive Load*, allerdings dient es der Automatisierung der Schemata und trägt letztendlich zum Lernen bei. Ziel jeder Instruktion muss es sein, möglichst viel kognitive Kapazität für diese Kategorie frei zu halten, damit Lernen stattfinden kann. Das bedeutet, dass im Idealfall die *Extraneous Cognitive Load* so klein wie möglich gehalten wird, weil diese Kategorie der Hauptansatzpunkt für Veränderung der CL ist. Nur im Notfall kann auch die *Intrinsic Cognitive Load* verringert werden, allerdings zahlt man dafür den möglicherweise hohen Preis, dass ein Material oder ein Unterrichtsabschnitt

kontextlos und ungebunden dargestellt werden muss, um eine Überforderung zu vermeiden. Dieser Schritt könnte sogar in einer zeitweisen Abwendung von der wissenschaftlichen Arbeitsweise resultieren, da *problem solving* und problemorientiertes Arbeiten laut einiger Cognitive Load Forscher einen zu hohen kognitiven Anspruch hat (vgl. PAAS, RENKL & SWELLER 2003). Diese Gefahr lässt sich besonders bei bilinguaem Biologieunterricht erkennen, da die sprachliche Barriere, besonders während des bilingualen Anfangsunterrichts, eine zusätzliche kognitive Belastung darstellt. Somit würden die Ängste einiger Biologielehrer, der bilinguale Biologieunterricht könnte fachlich abflachen, bestätigt werden.

In der *Cognitive Load* Forschung wird weiterhin darauf hingewiesen, dass sich die drei Kategorien zwar additiv zueinander verhalten (s. Abb.1), dass ihre Beziehungen untereinander und zur gesamten *Cognitive Load* jedoch asymmetrisch sind. Die *Intrinsic Cognitive Load* stellt den Basiswert dar, die verbleibenden Kapazitäten können verwendet werden, um mit der *Extraneous* und der *Germane Cognitive Load* umzugehen.

These can work in tandem in that, for example, a reduction in extraneous cognitive load by using a more effective instructional design can free capacity for an increase in germane cognitive load. If learning is improved by an instructional design that reduces extraneous cognitive load, the improvement may have occurred because the additional working memory capacity freed by reduction in extraneous cognitive load has now been allocated to germane cognitive load (PAAS ET AL. 2003, S. 2).

Dabei muss man immer wieder im Gedächtnis behalten, dass die Generierung und Automatisierung von Schemata letztendlich die *Intrinsic Cognitive Load* verringert und somit ein spiralförmiges Lernen ermöglicht wird.

Wie oben bereits beschrieben, trifft die *Cognitive Load Theory* nur auf neuartige Informationen zu. Wenn jedoch bereits Wissen zu diesen Informationen vorliegt, sei es im eigenen Langzeitgedächtnis oder durch Instruktion einer dritten Person, die zur Lösung des Problems dienen, werden diese höchstwahrscheinlich vorrangig genutzt. Dieser Aspekt begründet das Vorhandensein von Lehrern in schulischer Bildung. Sinnvolle Instruktion kann dabei helfen, die Limitierungen des Arbeitsgedächtnisses zu umgehen oder aufzulösen. Andererseits kann fehlerhafte Instruktion aber auch einen gegenteiligen Einfluss haben. Gerade in einem

Unterricht, der eine so hohe Komplexität aufweist, wie bilingualer Biologieunterricht mit bioethischen Problemsituationen, muss darauf geachtet werden, dass die Instruktion durch den Lehrer so sinnvoll ist, dass das Arbeitsgedächtnis nicht überlastet wird und dass die *Cognitive Load* so niedrig wie möglich gehalten wird, um Lernprozesse zu gewähren. Im Folgenden werden einige Effekte dargestellt, die genau diese Anforderungen behandeln und aus Forschungen zur *Cognitive Load Theory* hervorgegangen sind.

2.2.2 Ausgewählte Effekte der *Cognitive Load Theory* auf den Lernprozess

In den jüngeren Forschungen zur CLT wurde der Fokus der Theorie ständig erweitert. Es wurden mehrere Effekte ermittelt, die einen Einfluss auf die Belastung des Arbeitsgedächtnisses haben. Einige dieser Effekte haben aufgrund der Bedingungen dieser Arbeit einen Einfluss.

Der *Split-Attention Effect* (CHANDLER & SWELLER 1992; SWELLER 2005) tritt auf, wenn ein Lernender mit verschiedenen Quellen oder Materialien gleichzeitig beschäftigt ist. „[It] occurs when attention must be split between multiple sources of visual information that are all essential for understanding“ (SWELLER 2005, S. 26f.). Der gleiche Effekt würde eintreten, wenn verschiedene Informationsquellen einen anderen sinnlichen Kanal betreffen. Neuere Forschungen zeigen hingegen, dass eine Verbindung von visuellen und auditiven Informationen nicht vom *Split-Attention-Effect* betroffen sind. SWELLER nennt als Beispiel eine Aufgabe, bei der ein Diagramm (erste visuelle Quelle) mit zugehörigen Aussagen (zweite visuelle Quelle) kombiniert wird. „The multiple sources must be mentally integrated before the instruction can be understood and the material learned“ (ibid., S. 27). Wird das Diagramm allerdings visuell präsentiert und die Statements dazu auditiv, dann findet eine Integration schneller und mit weniger kognitivem Aufwand statt. Angewendet auf den bilingualen Biologieunterricht ergibt sich aus dem *Split-Attention Effect* ein Grund für eine kognitive Mehrbelastung. Bilinguales Fachwissen und naturwissenschaftlich-sprachliches Wissen sind Informationen aus zwei Quellen, die zuerst integriert werden müssen. Dies ist vor allem beim Anfangsunterricht und zu Beginn neuer Themen der Fall, weil die beiden Wissensfelder integriert werden müssen. Mögliche Hilfestellungen, wie Vokabellisten dürften das Problem hierbei theoretisch nicht lösen, weil sie auch eine weitere visuell dargebotene Information

darstellen, der ein gewisser Teil der kognitiven Kapazität gewidmet werden muss. Dennoch sind Vokabellisten und Glossare eine willkommene Hilfestellung (vgl. Kap. 2.2.3).

Ein weiterer Effekt, der in die Planung und Gestaltung von Material für den bilingualen Biologieunterricht einbezogen werden sollte, ist der *Expertise Reversal Effect*. Dieser beschreibt die Lernphasen, in denen Zusatzinformationen, Hilfestellungen und Materialerleichterungen zuerst insignifikant und dann kontraproduktiv werden. In späten Lernphasen, in denen schon zahlreiche Schemata zu der Thematik aufgebaut und automatisiert wurden, werden Zusatzinformationen überflüssig und behindern den Lernprozess. „Information that is essential for novices becomes redundant for more expert learners” (SWELLER 2005, S. 27). Es ist folglich für Lehrer im bilingualen Unterricht wichtig zu erkennen, in welcher Lernphase sich die Schülerinnen und Schüler befinden und in welchem Maße sie Unterstützung und Hilfsmechanismen benötigen. Die Behinderung von Schülerinnen und Schülern in der späten Lernphase durch eine Vokabelliste lässt sich dadurch erklären, dass sie trotz ihres Wissens an Fachbegriffen und Kontexten aus empfundener Unsicherheit auf die Liste zurückgreifen und dadurch wichtige kognitive Ressourcen verschwenden.

Während der *Split Attention Effect* und der *Expertise Reversal Effect* die *Extraneous Cognitive Load* beeinflussen, wird die *Intrinsic Cognitive Load* durch die *Element Interactivity* (SWELLER 1994; SWELLER 2006) bestimmt. Die *Element Interactivity* beschreibt die inhaltliche Verbindung der Aspekte, die gelernt werden. „When the elements of a task can be learned in isolation, they will be described as having low element interactivity. The level of element interactivity or connectedness refers to the extent to which elements of a task can be meaningfully learned without having to learn the relations between any other elements” (SWELLER 1994, S. 304). Für den Biologieunterricht bedeutet das, dass dessen Themen in der Regel eine hohe *Element Interactivity* haben. Das Thema Evolution beispielsweise muss immer wieder mit einbezogen werden, wenn genetische oder ökologische Themen behandelt werden. Gleichzeitig gibt es aber auch in der Biologie Felder ohne große Verbindungen zu anderen Elementen. Ein Schüler kann beispielsweise die Größe eines Erythrozyten lernen, ohne Verbindungen zu anderen Aspekten zu benötigen. Die *Element Interactivity* ist somit der ausschlaggebende Faktor für die Höhe der *Intrinsic*

Cognitive Load (vgl. SWELLER 2006). *Element Interactivity* ist allerdings nicht mit Komplexität oder Schwierigkeitsgrad einer Aufgabe gleichzusetzen. Auch Aufgaben oder Materialien mit einer geringen *Element Interactivity* können sehr schwer zu lernen sein. Wenn man beispielsweise Vokabeln einer Fremdsprache lernt, weisen diese Vokabeln oftmals nur eine Interaktivität mit ihrem muttersprachlichen Gegenüber auf, dennoch sind sie schwer zu lernen, wenn die Anzahl der Vokabelpaare ausreichend groß ist.

2.2.3 Anpassung der Theorie an den bilingualen Biologieunterricht

Die *Cognitive Load Theory* befasst sich mit der Ermittlung der Belastung, die von einer Lernumgebung, z.B. einer Aufgabe ausgeht. Die Anforderungen der KMK und der Kultusministerien widersprechen sich in einigen Punkten allerdings mit den Befunden der CLT. Gerade offene, problemlösungsbasierte Aufgaben sollen vermehrt eingesetzt werden. Folglich muss die Theorie an die Unterrichtswirklichkeit angepasst werden.

Die *Cognitive Load Theory* und die Forschungsergebnisse der letzten Jahre „almost uniformly support[...] direct, strong instructional guidance rather than constructivist-based minimal guidance during the instruction of novice to intermediate learners. Even for students with considerable prior knowledge, strong guidance while learning is most often found to be equally effective as unguided approaches” (KIRSCHNER ET AL. 2006, S. 83f.). Diese Befunde scheinen auf den ersten Blick den Unterrichtstheorien im Biologieunterricht zu widersprechen, die auf holistisches Lernen mit möglichst hoher Schülerzentrierung und Selbstständigkeit ausgelegt sind. Es stellt sich allerdings die Frage, inwieweit die Schülerinnen und Schüler, die in Deutschland BU im Fach Biologie belegen, noch als Anfänger zu bezeichnen sind. Die zahlreichen terminologischen Ähnlichkeiten in der englischen und deutschen Fachsprache im Fach Biologie lassen eine gewisse Vorbildung voraussetzen, die selbst Schülerinnen und Schüler haben, die gerade mit dem bilingualen Biologieunterricht angefangen haben. Gleichzeitig hat jahrelanger Englischunterricht und gegebenenfalls BU in anderen Sachfächern die Schülerinnen und Schüler soweit mit fachlichem Arbeiten unter Benutzung der englischen Sprache als Medium vertraut gemacht, dass von Anfängern eigentlich kaum noch die Rede

sein kann. Wirklich neuartig sind allerhöchstens die biologischen Inhalte und die damit verbundene neuartige Terminologie.

Wenn die Schülerinnen und Schüler nun im Testheft mit einem neuartigen Material konfrontiert sind, so ziehen sie verschiedene Schemata zur Bearbeitung des Materials heran. RENKL UND ATKINSON, die den Begriff Schemata auch als Kompetenz oder Fähigkeit verstehen postulieren, dass „learners may be simultaneously in different stages with respect to different parts of a skill“ (RENKL & ATKINSON 2003, S. 16). Die beiden Autoren unterscheiden dabei unter drei verschiedenen Phasen, nämlich einer frühen Phase, in der ein generelles Verständnis des Materials oder des Problems abläuft, ohne dass ein Schemata oder eine Kompetenz angewendet wird. In der mittleren Phase lernen die Schülerinnen und Schüler dann, wie ein Problem gelöst wird, wobei in der Regel auf einfache Probleme mit Lösungsanregungen zurückgegriffen wird. In der späten Phase werden die Geschwindigkeit und die Richtigkeit der Problemlösung immer weiter erhöht. Hierbei sollte schließlich echtes, selbstständiges Problemlösen durch die Schülerinnen und Schüler ausgeführt werden, um ein wirkliches Lernen und Verstehen zu gewährleisten (vgl. SWELLER 1994; RENKL & ATKINSON 2003). „In the late stage of skill acquisition, the major goal to be achieved is to heighten speed and accuracy. [...] When automaticity is the goal, self-explanations are not very helpful. Actually solving problems or part of them is the major path by which speed and accuracy can be enhanced“ (RENKL & ATKINSON 2003, S. 18). Dieser Aspekt ist bereits als *Expertise Reversal Effect* weiter oben beschrieben worden.

In welcher Phase sich die Schülerinnen und Schüler befinden, muss jeder Lehrer selber erkennen. Dabei muss er sich bewusst sein, dass verschiedene Phasen erreicht und abgeschlossen sein können. Hier ist eine Individualisierung der unterstützenden Maßnahmen angebracht, um einen gleichmäßigen Übergang zwischen den Phasen zu gewährleisten. Um diesen Übergang zu erreichen, schlagen VAN MERRIËNBOER ET AL. den Aufbau eines Gerüsts von Hilfestellungen, sogenanntes *Scaffolding*, vor. Wie beim Hausbau wird ein Gerüst angebracht, welches das neue Gebäude stützt und nach und nach zurückgebaut wird. „Scaffolding explicitly pertains to a combination of performance support and fading. Initially, the support enables a learner to achieve a goal or action not achievable without that support. When the learner achieves the desired goal, support gradually diminishes until it is no longer needed“ (VAN

MERRIËNBOER ET AL. 2003, S. 5). Es muss bei der Stellung von Hilfen allerdings darauf geachtet werden, dass die Hilfen möglichst in das Material integriert werden, um einen *Split Attention Effect* zu vermeiden. Die oben bereits beschriebene Vokabelliste ist in diesem Kontext weniger gut geeignet, da weniger integriert als Vokabelanmerkungen und Fußnoten im Material selber. Van Merriënboer und seine Kollegen schlagen aber noch subtilere Hilfsmöglichkeiten vor. Zu Beginn des Lernprozesses sollen die Aufgaben selber soweit heruntergebrochen werden, dass sie immer wieder auf einen gleichen Ablauf der Lösung hinauslaufen, der gelernt und angewendet wird. Diese sehr engen Aufgabenstellungen des Typs *Worked-Out Examples* sollte dann jedoch nach und nach abgelöst werden. „Different types of learning tasks thus range from tasks with high built-in support (worked-out examples), via tasks with an intermediate level of support (completion, goal-free, and reverse tasks), to conventional tasks without support” (VAN MERRIËNBOER ET AL. 2003, S. 8). Richtig angewendet können die Hilfestellungen die *Cognitive Load* empfindlich verringern und eine kognitive Verarbeitung ermöglichen, jedoch muss bedacht werden, dass die Verringerung der *Cognitive Load* nicht das Ziel der Hilfestellung sein soll. Das Ziel ist Hilfe zum Lernen.

Zusammenfassend lässt sich sagen, dass beim Einsatz von Hilfestellungen immer der *Split-Attention Effect* sowie der *Expertise Reversal Effect* bedacht werden müssen, um sinnvolle und nützliche Hilfestellungen zu entwickeln.

2.3 Der Kompetenzbereich Bewerten

2.3.1 Bewertungskompetenz im Zusammenhang mit den Bildungsstandards

Zusammen mit den anderen drei Kompetenzbereichen Fachwissen, Erkenntnisgewinnung und Kommunikation bildet der Bereich Bewerten die vier Kompetenzbereiche der nationalen Bildungsstandards für den Mittleren Schulabschluss im Fach Biologie. Man könnte nun fragen, warum ein solcher Kompetenzbereich formuliert wurde. CORINNA HÖBLE (Höble 2007a) begründet diesen mit den Grundrechten und Werten der Verfassung der Bundesrepublik Deutschland. Unser demokratisches System hat die Aufgabe, diese Grundrechte durchzusetzen, ist bei der Aufrechterhaltung dieser allerdings auf moralbildende Institutionen angewiesen, weil jeder Bürger die Werte erst kennenlernen und

verstehen muss, um vollständig am gesellschaftlichen Leben teilnehmen zu können. Somit übernimmt die Schule gesellschaftliche Relevanz und ist nach HÖBLE eine lebendige ethische Kraft. „Das Ziel kann nur eine verantwortliche Mündigkeit sein, nicht einfach die Anpassung an die Gesellschaft“ (HÖBLE 2007a, S. 197). Dies bedeutet, dass die Schule genau der Ort sein kann, an dem sich junge Menschen mit Werten und Normen auseinandersetzen. Gerade Gymnasiasten, denen man prophezeit, zukünftige Entscheidungsträger in unserer Gesellschaft zu sein (vgl. LIND 2006), aber auch allen anderen Schülerinnen und Schülern soll durch die Einführung einer unterrichtsverbindlichen Bewertungskompetenz diese Auseinandersetzung ermöglicht werden. Dafür reicht ein Behandeln von Aufgaben und Szenarien aus den Bereichen Bioethik und Nachhaltige Entwicklung natürlich nicht aus. „Eine erfolgreiche Bearbeitung von komplexen Problemen und Entscheidungssituationen erfordert einen naturwissenschaftlichen Unterricht, der Lernprozesse auf die Interaktion von Schülerinnen und Schülern ausrichtet und Diskussions- und Aushandlungsprozesse fördert“ (EGGERT ET AL. 2010, S. 300). Natürlich sind Schülerinnen und Schüler auch ohne Biologieunterricht in der Lage, Entscheidungen zu treffen, die sie auch begründen können. Sogar ethisch-moralische Entscheidungen werden getroffen, allerdings wurde empirisch belegt, dass die Entscheidung oftmals intuitiv getroffen wird und dass diese intuitiven Entscheidungen dann im Nachhinein durch Rechtfertigungen zu retten versucht werden. Diese Art von Entscheidung und Entscheidungsbegründung ist jedoch nicht mit jenen zu vergleichen, die den Zielen der Bildungspolitik entsprechen. „Aus bildungswissenschaftlicher und bildungspolitischer Sicht ist der Anspruch an Bewertungskompetenz aber, Schüler(innen) zu systematischen und begründeten Entscheidungen in komplexen Situationen [...] zu befähigen“ (EGGERT & BÖGEHOLZ 2006, S. 181). Dieses Bewusstsein findet sich dann auch in fast allen vorgeschlagenen Modellen zur Urteilsbildung wieder. Von den zehn Modellen, die BÖGEHOLZ ET AL. (2004) zusammengestellt haben, erfolgt die Entscheidungs- bzw. Beurteilungssituation im letzten oder in einem der letzten Schritte und wird dann nur noch durch eine abschließende Reflexion der Folgen der Entscheidung ergänzt.

In der Biologiedidaktik wurden mehrere Modelle zur Bewertungskompetenz und zugehörigen Teilkompetenzen veröffentlicht. Diese Modelle werden Kompetenzstrukturmodelle genannt, weil sie darlegen, aus welchen Teilkompetenzen

oder Dimensionen die Bewertungskompetenz besteht. Sie stellen somit eine Liste von Voraussetzungen, und Anforderungen dar, die den Schülerinnen und Schülern vermittelt und von diesen eingeübt werden müssen, damit das Ziel der Bewertungskompetenz erreicht werden kann. Dieses Ziel kann definiert werden als „die Fähigkeit, die ethische Relevanz naturwissenschaftlicher Themen wahrzunehmen, damit verbundenen Werte zu erkennen und abzuwägen sowie ein reflektiertes und begründetes Urteil zu fällen“ (MITTELSTEN SCHEID & HÖBLE 2007, S. 88).

Die prominentesten Kompetenzstrukturmodelle sind das Göttinger Modell von SABINE EGGERT und SUSANNE BÖGEHOLZ und das Oldenburger Modell von KATJA REITSCHERT. Das Göttinger Modell geht von vier Säulen aus, von der die Bewertungskompetenz gestützt wird. Die erste Säule behandelt „Kennen und Verstehen von Nachhaltiger Entwicklung“, es folgen „Kennen und Verstehen von Werten und Normen“, „Generieren und Reflektieren von Sachinformationen“ und „Bewerten, Entscheiden und Reflektieren“. Die erste Säule ist „für andere Kontexte, wie z.B. der Medizin- oder Bioethik, durch das jeweilige in dieser Domäne relevante Wissen auszutauschen“ (EGGERT & BÖGEHOLZ 2006, S. 189). In der Regel konzentriert sich das Göttinger Modell vorwiegend auf die Thematik der nachhaltigen Entwicklung.

Als theoretische Grundlage dieser Arbeit und des zugrundeliegenden Fragenkatalogs wurde deshalb das Oldenburger Modell ausgewählt, dass einerseits nicht auf eine Thematik festgelegt ist und andererseits eine präzisere Unterteilung der Teilkompetenzen aufweist (vgl. Abb.2). Die vierte Säule des Göttinger Modells ist beim Oldenburger Modell beispielsweise in drei Teilkompetenzen unterteilt.

Die Teilkompetenzen „Wahrnehmen und Bewusstmachen der eigenen Einstellung“ und Wahrnehmen und Bewusstmachen moralischer Relevanz“ sind die beiden Teilkompetenzen, die die Ausgangsstellung der Bewertung darstellen. Beide erfordern eine kritische Reflexion der eigenen Einstellung, dem Ursprung des eigenen Wertsystems und eine Sensibilität für moralisch-ethische Situationen.

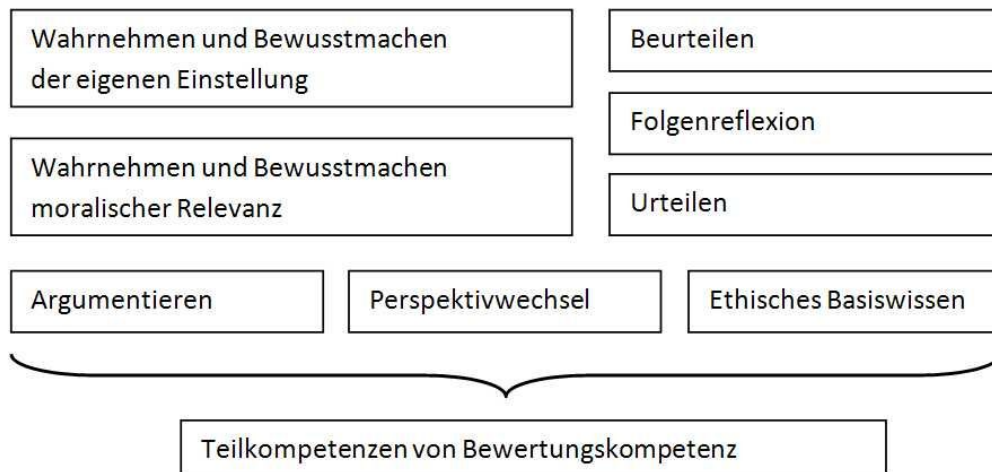


Abbildung 2: Teilkompetenzen von Bewertungskompetenz (nach REITSCHERT ET AL. 2007, aus MITTELSTEN SCHEID & HÖBLE 2007).

Die persönliche Prägung ist maßgeblich dafür entscheidend, eine moralische Relevanz überhaupt wahrzunehmen und hängt von der Erziehung, der Zugehörigkeit zu einer Kultur oder Religion, aber auch vom Biologieunterricht ab.

Die beiden Teilaspekte „Beurteilen“ und „Ethisches Basiswissen“ stehen sich gewissermaßen gegenüber, da bei ersterem eine deskriptive Analyse der Situation erforderlich ist, bei der Fachwissen essentiell ist, während bei letzterem ein ethisches Grundwissen erforderlich ist, das der Schüler haben muss, um bioethischem Unterricht folgen und sinnstiftend daran teilnehmen zu können.

Da im Biologieunterricht in den Situationen zwar Entscheidungen getroffen werden, diese jedoch nur im seltensten Fall umgesetzt werden (können), ist die Teilkompetenz der „Folgenreflexion“ ein besonders essentieller Teil der Bewertungskompetenz. Auch hier strahlt das Fachwissen stark mit ein. Die Schülerinnen und Schüler müssen zwischen definitiven und möglichen Folgen abwägen und eruieren, ob die von ihnen getroffene hypothetische Entscheidung Folgen hervorbringt, die ein neues Dilemma erzeugen oder überhaupt mit den Werten des Entscheidungsträgers vereinbar sind.

„Perspektivwechsel“ und „Argumentieren“ sind zwei Teilkompetenzen, „die kaum aus den anderen Dimensionen herauszulösen sind [und sich durch] den gesamten Prozess einer ethischen Urteilsbildung [ziehen]“ (REITSCHERT ET AL. 2007, S. 43). Für ein begründetes Urteil ist es unbedingt notwendig, alle betroffenen Parteien und deren Sichtweisen auf das Problem oder das Dilemma zu betrachten. Dieser Prozess

erfordert ein gewisses Maß an Empathie, Verständnis, Toleranz und Sensibilität. Sie steht auch der Bewusstmachung der eigenen Einstellung gegenüber und schließt somit den Kreis.

Die letzte Teilkompetenz ist das „Urteilen“ selber, bei dem eine Entscheidung getroffen werden muss. Dieses Urteil soll gut begründet sein, ist allerdings offen für Veränderungen, wenn neue Folgen oder Sichtweisen einbezogen werden. Das Urteil stellt im schulischen Kontext oftmals das Endprodukt des Bewertungsprozesses dar, weil die getroffene Entscheidung nicht umgesetzt werden kann.

Die Formulierung einer Kompetenz macht nur Sinn, wenn sie mit Inhalten gefüllt ist. Im Falle der Kompetenzbereiche im Fach Biologie für den Mittleren Schulabschluss wurden folglich sieben Regelstandards formuliert, die an den Schulen implementiert werden sollen (vgl. KMK 2005). Um die Bewertungskompetenz der Schülerinnen und Schüler konkreter bewerten zu können wurden weiterhin verschiedenen Niveaus der Bewertungsteilkompetenzen erstellt, die die Leistung der Schülerinnen und Schüler im Bereich einer Teilkompetenz bewertet. MITTELSTEN SCHEID & HÖBLE (2007) sowie REITSCHERT ET AL. (2007) formulieren Niveaustufen für die Teilkompetenzen von Bewertungskompetenz des Oldenburger Modells. Diese Niveaustufen liegen den bioethischen Aufgaben des Testheftes zugrunde.

Trotz aller theoretischer Vorarbeit, die seit der Einführung der Bildungsstandards geleistet wurde, zeigt sowohl die empirische Datenlage, als auch die Implementierung im täglichen Biologieunterricht bestehende Defizite. So hat die Bewertungskompetenz bislang keinen Eingang in die Vorgaben des Landesabitur Hessens gefunden, wie die Lehrkräfte der Biologiekurse, die an der Befragung für diese Arbeit teilgenommen haben, berichteten. In empirischen Untersuchungen, die zur bewertenden Leistung von Schülerinnen und Schülern durchgeführt wurden, konnte gezeigt werden, „dass Schüler(innen) Schwierigkeiten im Umgang mit einer systematischen Vorgehensweise in Entscheidungsprozessen haben und über kaum Metastrategie- bzw. Bewertungsstrukturwissen verfügen“ (EGGERT & BÖGEHOLZ 2006, S. 191). Die bewertungs- und urteilsrelevanten Folgen sind absehbar: ohne dieses Wissen können die Schülerinnen und Schüler nur im Stadium der intuitiven Entscheidung verbleiben, dass bereits oben diskutiert wurde.

2.3.2 Bewerten im bilingualen Biologieunterricht

Wie bereits in Kapitel 2.1 dargestellt, liegen der Lehrplan und die Voraussetzungen des Sachfaches diesem auch im Kontext des bilingualen Unterrichts zu Grunde. Wenn also ethisch-moralisches Bewerten als verbindliche Kompetenz für das Fach Biologie eingefordert wird, so müssen auch die Schülerinnen und Schüler in bilingualen Kursen dieser Auflage Folge leisten. Betrachtet man aber den theoretischen Rahmen der Cognitive Load Theory parallel zu dieser bildungspolitischen Forderung, kommen Fragen auf, ob und wie dieses Ziel erreicht werden kann.

Ein Problem bei der Umsetzung ist sicherlich ein Mangel an geeigneten, muttersprachlichen Materialien für den BU. Authentisches Material ist artifiziell didaktisch aufbereitetem, zumindest bei der Behandlung der Themen Bioethik und nachhaltige Entwicklung, vorzuziehen. Das Konzept der ethisch-moralischen Entscheidung ist im angelsächsischen Raum für den schulischen Unterricht allerdings weitgehend unbekannt, weshalb bei dem für diese Arbeit verwendeten Testheft auch auf deutsche Aufgabenbeispiele zurückgegriffen werden musste. Im angelsächsischen Sprachraum hat sich zwar der Begriff der *decision making competence* etabliert (vgl. EGGERT & BÖGEHOLZ 2006), allerdings haben Aufgaben, die diese Kompetenz fördern sollen, einen anderen Fokus. Für den bilingualen Biologieunterricht müssen daher dringend englischsprachige Materialien geschaffen werden, die diesen Missstand beheben.

3 Fragestellung

3.1 Formulierung der Fragestellung

Die Forschung zu bilingualem Unterricht fokussiert sich bislang hauptsächlich darauf, ob bilingual unterrichtete Schülerinnen und Schüler die gleichen Lernerfolge aufweisen, wie ihre auf Deutsch unterrichteten Mitschüler. Diese Gleichheit der Lernerfolge wird als Minimalziel des bilingualen Unterrichts beschrieben. Zwar bildet diese Fragestellung auch für diese Arbeit einen Ansatz, allerdings liegt der Hauptfokus hier auf der Frage, ob ein Zusammenhang zwischen der kognitiven Leistungsfähigkeit und der Bewertungskompetenz von Schülerinnen und Schülern besteht. Die ermittelten Zusammenhänge und die daraus gezogenen Schlussfolgerungen können schließlich wieder auf das Minimalziel bezogen werden.

Zusätzlich zu den Befunden hinsichtlich des bilingualen Unterrichts kann diese Studie noch weitere wertvolle Einblicke in einen anderen Forschungsbereich gewähren. Die Kompetenzforschung und Kompetenzerwerbsforschung befindet sich für deutschsprachigen Biologieunterricht noch in den Anfängen, weshalb diese Arbeit auch einen Beitrag zur allgemeinen Aufklärung beitragen kann. Unabhängig von der Unterrichtssprache der Schülerinnen und Schüler, die an der Studie teilnehmen, können Aussagen über die Ausprägung der Bewertungskompetenz getätigt werden.

Im Bezug auf bilingualen Unterricht in Verbindung mit Bildungsstandards stellt sich die Forschungslage ambivalent dar. Die Befunde, dass bilinguale Schüler eine höhere Kompetenz sowie einen höheren Kompetenzzuwachs aufweisen, wie dies für sprachliche Fähigkeiten empirisch belegt wurde, können auf die sachfachspezifischen Kompetenzen noch nicht übertragen werden. Zwar weisen einige Studien darauf hin, dass Schülerinnen und Schüler, die bilingualen Biologieunterricht besuchen, leicht bessere oder zumindest keine schlechteren Leistungen abliefern (vgl. Osterhage 2009), allerdings stellt sich die Forschungslage insgesamt als unzureichend dar. Diesen wenigen empirischen Forschungsergebnissen, die einen Mehrwert des bilingualen Unterrichts attestieren, steht die *Cognitive Load Theory* gegenüber, die eine kognitive Mehr- bis Überbelastung der Schülerinnen und Schüler im bilingualen Biologieunterricht proklamiert. Insbesondere in komplexen Unterrichtsinhalten, wie sie im Kontext der Bewertungskompetenz üblich sind, müsste es, folgt man der *Cognitive Load Theory*, nur kognitiv leistungsfähigen Schülerinnen und Schülern gelingen, gute, d.h. vergleichbare Leistungen zu erbringen wie ihre auf Deutsch unterrichteten Mitschüler. Überträgt man die Befunde der Arbeit von der Fragestellung auf den Biologieunterricht, dann soll in dieser Arbeit geklärt werden, ob englischsprachiger Biologieunterricht die Zielsetzungen eines Naturwissenschaftsunterrichts angemessen erreichen kann. Hat die kognitive Leistungsfähigkeit der Schülerinnen und Schüler einen Einfluss auf deren Leistungen im bilingualen Biologieunterricht im Bezug zur Bewertungskompetenz bei komplexen Themenfeldern der Bioethik und der nachhaltigen Entwicklung? Dabei muss der Vergleich mit muttersprachlich erteiltem Biologieunterricht immer bedacht werden (vgl. KOCH/BÜNDER 2006).

Wie in Kapitel 2 dargestellt, bieten viele Schulen bilingualen Biologieunterricht an, ohne dass es eine fachdidaktische Grundlage dafür gibt. Bei diesen Schulen handelt es sich um solche, die vorrangig modularen bilingualen Unterricht anbieten. Bei diesen Schulen scheint die Gefahr der kognitiven Überlastung der Schülerinnen und Schüler besonders hoch zu sein, weil diese weniger an das Denken und Arbeiten in der englischen Sprache als Kommunikations- Transportmedium gewöhnt sind und die resultierende Umstellung wertvolle kognitive Ressourcen von Lernprozessen abzieht.

Im Zuge dieser Arbeit wurden daher Teilforschungsfragen formuliert, die der Grundfragestellung folgen, diese aber um einen Parameter erweitern. Dadurch sind sie in ihrer Gesamtheit der Beantwortung der Fragestellung zuträglich. Die Struktur dieser Unterfragestellungen schließt immer die Grundstruktur der untersuchten Korrelation von kognitiver Leistungsfähigkeit und Bewertungskompetenz bzw. den Leistungen der Schülerinnen und Schüler mit ein. Hinzugefügt werden allerdings die Parameter Alter/Jahrgangsstufe, Unterrichtsform monolingual-bilingual, sowie die Unterscheidung bilingualer Zug – bilinguales Modul. Die Fragestellungen lauten wie folgt:

- Gibt es einen Zusammenhang zwischen dem Alter bzw. der Jahrgangsstufe der Schülerinnen und Schüler, ihrer kognitiven Leistungsfähigkeit und ihrer Bewertungskompetenz in den Feldern Bioethik und nachhaltige Entwicklung?
- Gibt es einen Zusammenhang zwischen der Bewertungskompetenz von Schülerinnen und Schülern, ihrer kognitiven Leistungsfähigkeit und der von ihnen besuchten Unterrichtsform?
- Haben Schülerinnen und Schüler in bilingualen Zügen eine größere Bewertungskompetenz als Schülerinnen und Schüler, die modularen BU Biologie belegen?

3.2 Hypothesen

Zieht man die *Cognitive Load Theory* (vgl. Kap. 2.2) in die Forschungsfragen mit ein, so kann man die Hypothese aufstellen, dass Schülerinnen und Schüler mit einer hohen kognitiven Leistungsfähigkeit besser bei der Lösung der Testaufgaben

abschneiden. Folglich muss davon ausgegangen werden, dass bilingualer Biologieunterricht zumindest für den Kompetenzbereich Bewerten für Schüler mit einem relativ kleinen Arbeitsgedächtnis eine Überforderung darstellt, was sich in einer verringerten Bewertungskompetenz widerspiegelt. Die Cognitive Load der expliziten Aufgaben ist für Schülerinnen und Schüler mit geringer Leistungsfähigkeit höher. Diese Befunde könnten auf die Bewertungskompetenz und den bilingualen Biologieunterricht allgemein übertragen werden. Die Übertragbarkeit ist gerade deswegen gegeben, weil die Limitierungen des Arbeitsgedächtnisses im Kontext des Fragebogens eine sehr hohe Rolle einnimmt. Da die Schülerinnen und Schüler beim Ausfüllen des Fragebogens mit neuartigen Aspekten konfrontiert sind, besteht in der Regel noch kein Vorwissen, welches die Bedeutung des Arbeitsgedächtnisses abschwächen könnte.

Schüler der 12. Klasse bzw. des gymnasialen Bildungsganges G9 haben in dem zusätzlichen Schuljahr bereits gewisse Schemata generiert und automatisiert, weshalb sie sich direkt auf das Problem konzentrieren können.

„When a complex intellectual skill is first acquired, it may be usable only by devoting considerable cognitive effort to the process. With time and practice, the skill may become automatic to the point where it may require minimal thought for its operation. It is only then that intellectual performance can attain its full potential. Without automation, performance is slow, clumsy and prone to error. It is an essential mechanism of learning” (SWELLER 1994, S. 298).

Auf dieser Grundlage kann die Hypothese aufgestellt werden, dass Schülerinnen und Schüler einer höheren Klassenstufe im Testheft deutlich besser abschneiden müssten, weil die Aufgaben für sie weniger komplex sind. Da ihre Schemata soweit vorhanden und automatisiert sind, können sie einen Großteil ihrer Arbeitsgedächtniskapazität auf die Lösung der Aufgaben verwenden, während jüngere Schüler zumindest Anteile der Kapazität zur Bewältigung der Komplexität aufwenden müssen. Folgt man der Argumentation der *Cognitive Load Theory*, so wird die Limitierung des Kurzzeitgedächtnisses irrelevant, sobald Schemata oder Vorwissen aufgebaut und automatisiert wird (vgl. PAAS, RENKL & SWELLER 2004). Diese kognitive Entwicklung vorausgesetzt, müssten ältere Schüler, d.h. Schüler einer höheren Jahrgangsstufe bzw. einer längeren bisher absolvierten Schulkarriere, eine höhere Bewertungskompetenz haben als die jüngeren Teilnehmer. Diese Hypothese müsste

sowohl für die bilingual als auch für die auf Deutsch unterrichteten Schülerinnen und Schüler Gültigkeit haben.

Zwar sind Schülerinnen und Schüler durch die organisatorischen Bedingungen von bilingualem Biologieunterricht zumindest zu Beginn der Sekundarstufe I einer höheren Stundenzahl ausgesetzt, was zu einem vermehrten Kompetenzaufbau führen müsste. Andererseits stellen komplexe Unterrichtsinhalte wie Bioethik und nachhaltige Entwicklung eine höhere kognitive Belastung für sie dar als für ihre auf Deutsch unterrichteten Mitschüler. Folglich kann die Hypothese aufgestellt werden, dass bilingual unterrichtete Schülerinnen und Schüler eine niedrigere Bewertungskompetenz aufweisen, die auf der höheren kognitiven Belastung beruht.

Schülerinnen und Schüler, die einen bilingualen Zug besuchen, sollten beim Fragebogen deutlich besser abschneiden als Schüler, die Biologie nur als modularen BU besuchen. Die Exposition der Zug-Schüler zum CLIL-Konzept in anderen Fächern führt zu einer höheren Sprach- und Strategiekompetenz, die auf den bilingualen Biologieunterricht übertragen werden kann. KOCH und BÜNDER haben in einer Studie von 2006 nachgewiesen, dass die deutlichsten Leistungssteigerungen im bilingualen Biologieunterricht bei denjenigen Kursen zu verzeichnen sind, „deren Teilnehmer bereits Erfahrungen mit dem englischsprachigen, bilingualen Unterricht hatten“ (KOCH/BÜNDER 2006, S. 72).

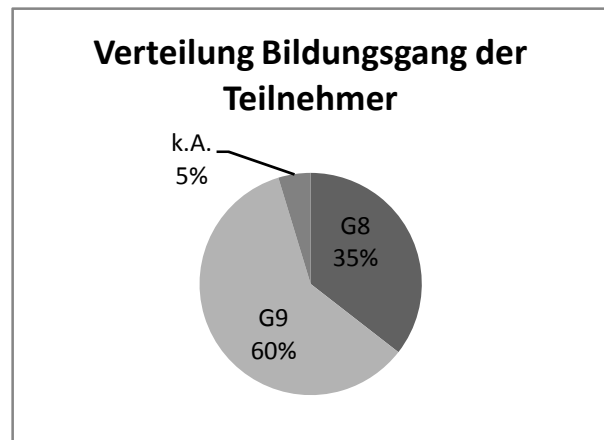
4 Methodisches Vorgehen

4.1 Die Stichprobe

Die Datenerhebung wurde in vier Kasseler Gymnasien in insgesamt neun Klassen durchgeführt. Die Stichprobe umfasst insgesamt 171 Schülerinnen und Schüler, von allen ausgefüllten Testheften konnten 169 für die Auswertung verwendet werden. Eine Schülerin konnte den Fragebogen aus gesundheitlichen Gründen nicht vollständig ausfüllen, eine weitere Teilnehmerin war eine Austauschschülerin. Zwei Schüler erschienen verspätet zum Unterricht, weshalb sie dem Kognitionstest nicht oder nur teilweise beiwohnen konnten. Zwar zeichnet sich Hessen durch eine hohe Zahl Gesamtschulen mit bilingualem Angebot aus, es konnte jedoch keine Gesamtschule im Raum Kassel gefunden werden, in der bilinguale Biologiekurse

wählbar sind. Der Umstand, dass Biologie nur selten als bilinguales Fach in der Mittelstufe angeboten wird, wird einer der Gründe für diesen Umstand sein.

Bei der Stichprobe ist auffällig, dass die meisten Schülerinnen und Schüler in der 12. Klasse sind. Der absolute Anteil der Teilnehmer aus Jahrgang 11 beträgt 39 (23%), während 130 Schülerinnen und Schüler (77%) die 12. Klasse besuchen. Das ist dem doppelten



Abiturjahrgang geschuldet. Dadurch sind Schüler, die zum Zeitpunkt der

Abbildung 3: Verteilung Bildungsgang (in Prozent) der Teilnehmer der Studie (n=169; k.A.= keine Angabe)

Befragung in der 12. Klasse sind, entweder Schüler des Bildungsganges G9 oder G8. Im Falle von G8 haben die Teilnehmer sozusagen die 11. Klasse ausgelassen, sind daher deutlich jünger als andere Kursmitglieder. Dieser Umstand wird in der Auswertung beachtet. Während mit acht Schülerinnen und Schülern 5% der Stichprobe keine Angabe zu ihrem Bildungsgang machten, entfallen 35% (60 Teilnehmer) auf den Bildungsgang G8 und 60% (101 Teilnehmer) auf den Bildungsgang G9 (vgl. Abb. 3). Von den 60 G8-Schülern belegen 32 bilingualen Unterricht in Biologie und 14 noch mindestens ein weiteres bilinguales Fach. Zwar sind die Gesamtzahlen der G8- und G9-Schüler an den Schulen nicht abgefragt worden und die Stichprobe ist mit n=32 sehr klein, allerdings scheint sich die Schulzeitverkürzung durch G8 nicht auf die Belegung bilingualer Fächer auszuwirken, was der Hypothese, dass BU kognitiv zu anspruchsvoll sei, teilweise widerspricht.

Das Alter der befragten Schülerinnen und Schüler schwankt zwischen 15 bis 21, wobei das Durchschnittsalter bei 17,32 Jahren liegt. Aus Abb. 4 kann die Mengenverteilung auf die jeweiligen Altersstufen entnommen werden.

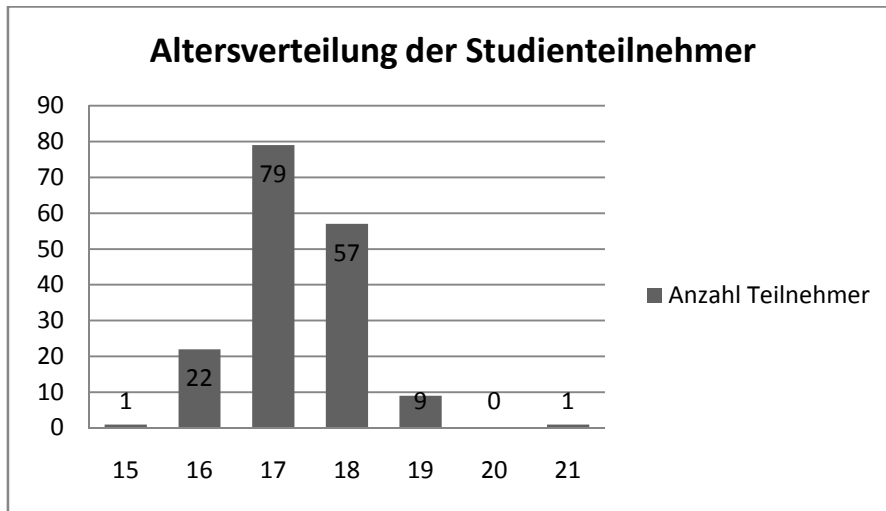


Abbildung 4: Absolute Altersverteilung der Studienteilnehmer (n=169)

Differenziert man die Stichprobe nochmals in Bio_{bili} und Bio_{deutsch}-Teilnehmer, so beträgt das Durchschnittsalter der Schülerinnen und Schüler in bili-Kursen 17,28 Jahre und bei den muttersprachlich unterrichteten Teilnehmern 17,39 Jahre.

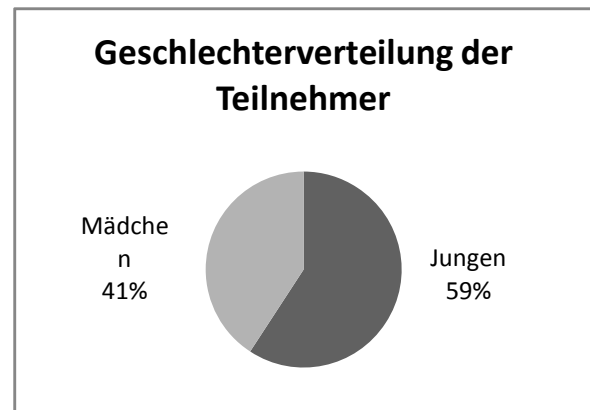


Abbildung 5: Geschlechterverteilung (in Prozent) der Teilnehmer der Studie (n=169)

Ein signifikanter Altersunterschied zwischen den beiden Gruppen ist folglich nicht zu erkennen.

Die Geschlechtsverteilung der Teilnehmer an der Studie ist leider unausgeglichen. Insgesamt haben 100 Schülerinnen und 69 Schüler an der Befragung teilgenommen, was einer prozentualen Verteilung von

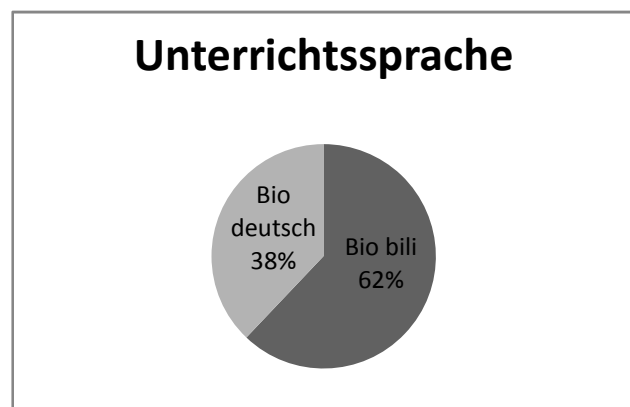


Abbildung 6: Verteilung der Unterrichtssprache (in Prozent) der Teilnehmer der Studie (n=169)

59,2% zu 40,8% entspricht (vgl. Abb. 5). Einer der Gründe für

dieses Ungleichgewicht ist sicherlich der Umstand, dass an einer der Schulen, an der drei Klassen befragt wurden, ein Schülerinnenanteil an der Gesamtschülerzahl von über 60% vorliegt. Die Lehrkräfte der anderen drei Schulen konnten keine entsprechenden Daten für ihre jeweilige Schule nachweisen. Da diese Studie jedoch einen Leistungsunterschied auf Grundlage des Geschlechts der Schüler nicht mit einbezieht, geht von dem Geschlechtsungleichgewicht keine allzu große Störung aus.

Auch bei der Verteilung der Stichprobe auf die beiden Unterrichtssprachen, bilingualer Biologieunterricht und Biologieunterricht auf Deutsch, gibt es ein Ungleichgewicht, das sogar noch größer ausfällt als bei der Geschlechtsverteilung. Insgesamt haben 104 Schülerinnen und Schüler, die bilingualen Biologieunterricht belegen, an der Studie teilgenommen. Demgegenüber stehen nur 64 muttersprachlich unterrichtete Schülerinnen und Schüler (vgl. Abb. 6). Dies resultiert in einem Verhältnis von 62,1% (Bio bili) zu 37,9% (Bio deutsch).

Die bisherigen Leistungen in den Fächern Biologie, sowie Englisch, Deutsch und

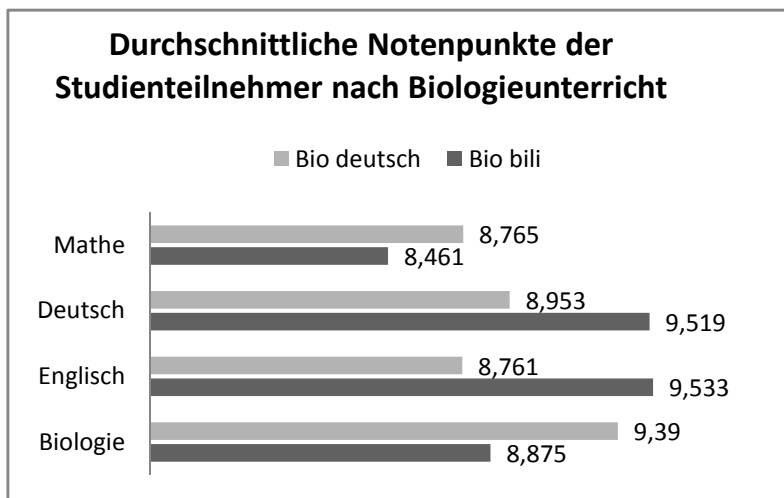


Abbildung 7: Durchschnittliche Notenpunkte der Studienteilnehmer nach Biologieunterricht; 15-Punkte-Skala (n=169)

Mathe wurden ebenfalls mit den persönlichen Daten abgefragt, um festzustellen, ob es schon vor dem Fragebogen signifikante Leistungsunterschiede gab. Dazu wurden die jeweils letzten Notenpunkte der oben genannten Fächer

abgefragt. Abbildung 7 zeigt die durchschnittlichen Notenpunkte. Wie zu erwarten war (vgl. BOHN 2008, FEHLING 2009, KRECHEL 2008), weisen die Teilnehmer aus den bilingualen Biologiekursen eine deutlich höhere durchschnittliche Englischnote auf, sie ist mit 9,533 Punkten um fast 0,8 Notenpunkte besser als bei Schülerinnen und Schülern, die deutschsprachigen Biologieunterricht belegten. Bei den Biologienoten ist dieses Verhältnis umgekehrt. Hier haben die deutschsprachig unterrichteten Teilnehmer einen höheren Durchschnitt (9,39), während jener der bili-Teilnehmer mit 8,875 um ca. 0,5 Notenpunkte niedriger liegt. Dieser Unterschied

unterstützt indirekt die These, dass Schülerinnen und Schüler, die bilingualen Biologieunterricht belegen, zwar eine höhere Sprachleistung, dafür aber ein geringeres Fachwissen erreichen können. Diese Beobachtung muss jedoch noch durch die Auswertung des Testheftes bestätigt werden, um hier eine signifikante Aussage machen zu können. Interessant ist allerdings, dass die bili-Teilnehmer auch im Unterrichtsfach Deutsch einen um fast 0,5 Notenpunkte höheren Notendurchschnitt aufweisen, als ihre deutschsprachig unterrichteten Mitschüler. Die Mathematiknoten unterscheiden sich bei den beiden Testgruppen nur um 0,3 Notenpunkte zu Gunsten der deutschsprachigen Stichprobe. Die mathematischen Kompetenzen haben keinen direkten Einfluss auf diese Untersuchung, können allerdings beim Kognitionstest zum Tragen kommen und tragen zu einem allgemeinen Überblick hinsichtlich einer naturwissenschaftlichen Begabung eines Schülers bei.

Für die dritte Unterfragestellung werden Schülerinnen und Schüler verglichen, die BU als bilingualen Zug oder als bilinguales Modul belegen. Allerdings bietet nur eine Schule einen bilingualen Zweig an. Daraus resultiert, dass die Stichprobe dieser Schüler im Vergleich zur Gesamtstichprobe unregelmäßig klein ausfällt ($n_{\text{anderer BU}} = 39$). Daher werden zu dieser Gruppe bei der Auswertung auch alle diejenigen Schüler gezählt, die Biologie zwar nur modular bilingual, zusätzlich aber noch mindestens ein anderes Fach bilingual belegen. Dadurch erhöht sich die Stichprobe auf $n=51$, was gemessen an den 104 Schülerinnen und Schülern, die bilingualen Biologieunterricht besuchen, 49,1% entspricht. Zwar belegen noch fünf weitere Teilnehmer anderen bilingualen Unterricht, allerdings werden sie nicht in diese Berechnung einbezogen, weil sie bilingualen Biologieunterricht nicht belegen.

4.2 Die Messinstrumente

Zur Durchführung der empirischen Untersuchung wurde ein Fragebogen erstellt, der aus drei Teilen bestand. Der erste Teil stellte eine kurze Erhebung der kognitiven Leistungsfähigkeit der teilnehmenden Schülerinnen und Schülern dar, den zweiten Teil bildeten sechs Testaufgaben zum Kompetenzbereich Bewerten und im dritten Teil wurden das subjektive Überforderungsempfinden der Schülerinnen und Schüler abgefragt.

Der Kognitionstest wurde in Anlehnung an MILLERS „magical number seven, plus or minus two“ (1956) konzipiert. Wie in den Ausführungen zur *Cognitive Load Theory* dargestellt, weist das Arbeitsgedächtnis die entscheidende Limitierung für das Lernen auf. Der Mensch ist nur in der Lage ungefähr sieben *Chunks* gleichzeitig aufzunehmen und zu verarbeiten. Eben diese von MILLER beschriebene Limitierung bildet den Ausgangspunkt des ersten Testanteils. Es wurde bewusst die Entscheidung getroffen, die kognitive Leistungsfähigkeit zu testen, und eben nicht die Cognitive Load der Aufgaben, die verwendet wurden. Zwar sind diese Messungen technisch durchaus möglich, übersteigen jedoch im Umfang dieser Arbeit die Mittel oder sind für die Zielsetzung der Arbeit nicht brauchbar. Zum Beispiel geben BRÜNKEN ET AL. einen Überblick über die momentan eingesetzten Testinstrumente. In der internationalen Forschung werden immer häufiger sogenannte *neuroimaging techniques* angewendet, bei denen mithilfe von Positronen-Emissions-Tomographie oder mit Funktioneller Magnetresonanztomographie direkt überprüft wird, welche Gehirnregionen bei der Bearbeitung von Aufgaben wie stark aktiviert werden. Eine weitere Möglichkeit ist die Messung der Pupillenerweiterung und Herzrhythmus beim Bearbeiten mit Aufgaben (vgl. BRÜNKEN ET AL. 2003). YOUSOOF ET AL. schlagen hingegen eine spontane Bewertung der *Cognitive Load* einer Aufgabe beim Lösungsvorgang vor (YOUSOOF ET AL. 2007).

Nach eingehender Sichtung von Tests zur kognitiven Leistungsfähigkeit, wurden zwei mögliche Testinstrumente isoliert, die sich in Aufbau und Durchführung stark ähnelten. Beide Testansätze machen sich die *Cognitive Load Theory* zunutze, da sie mit neuartigen Informationen arbeiten. Folglich wird der generelle Aufbau des Tests beschrieben. „New information such as new combinations of numbers or letters can only be stored for brief periods with severe limitations on the amount of such information that can be dealt with“ (KIRSCHNER ET AL. 2006, S. 77). In einem Ansatz bekommen die Teilnehmer einige kurze Sätze vorgelesen und müssen im Anschluss jeweils das letzte Wort jedes Satzes auflisten. „[S]ubjects might [...] hear ‚The sailor sold the parrot. The vicar opened the book‘. They should then respond ‚parrot, book‘. The test typically starts with two sentences and increases to a point at which subjects are no longer able to recall all terminal words. This point is designated the subject’s working memory span“ (BADDELEY 1992, S. 557). Bei einer Durchführung dieses

Testes wurden Werte ermittelt, die eine erstaunlich hohe Übereinstimmung mit MILLERS Beschreibungen aufweist.

Letztendlich fiel die Entscheidung aber gegen diesen Ansatz, weil er sprachlich begabtere Schülerinnen und Schüler bevorteilen könnte. Stattdessen wurde ein computergestützter Ansatz ausgewählt, der nach der *Digit-Span-Procedure* funktioniert und der entsprechend für den Schulgebrauch umgewandelt wurde (vgl. BRÄKLING 2007, MILLER 1956). Hierbei bekommen die Teilnehmer eine Ziffernfolge gezeigt, die sie sich merken sollen, ohne sie zu notieren. Nach der Ziffernfolge sind die Teilnehmer einer Ablenkung ausgesetzt bei der sie sieben leichte Additionsaufgaben lösen sollen. Dieses Testformat soll ebenso wie der oben beschriebene Test das Limit des individuellen Arbeitsgedächtnisses ermitteln. Im Anschluss daran sollen sie die Ziffernfolge aus dem Gedächtnis notieren. „If the digit-span procedure depends on the short-term store, with the number of digits retained determined by the capacity of the store, then it should be possible to interfere systematically with the operation of the working memory system by requiring the subject to remember digits while performing other cognitive tasks” (BADDELEY 1992, S. 556). Die Additionsaufgaben stellen dabei die andere kognitive Aufgabe dar, die zwischen der visuellen Aufnahme der Ziffernfolge und der Wiedergabe dieser stattfindet. Auch dieser Ansatz geht auf Millers Beobachtungen zurück. Der Vorteil dieses Instruments ist, dass ihn viele Teilnehmer simultan durchführen können, was bei vielen der oben beschriebenen Cognitive Load Tests nicht möglich wäre. Bei der Durchführung der Befragung werden die Ziffernfolge und die Additionsaufgaben digital an eine Wand projiziert oder bei unzureichender technischer Ausrüstung der Schule auf Karten hochgehalten. Die Anpassung des ursprünglich computerbasierten Tests auf ein *Paper-and-Pencil*-Format ist ebenfalls der Einsetzbarkeit im Schulkontext geschuldet. Die auftretenden Nachteile, vor allem die Anfälligkeit des Instruments für Betrug durch frühzeitiges Notieren der Ziffernfolge ist bedauerlich, allerdings überwiegen die Vorteile. Besondere Aufmerksamkeit bei der Testaufsicht ist erforderlich.

Der kognitive Leistungstest wurde in einem Prätest untersucht. Dabei wurden Ziffernfolgen von vier, fünf, sechs und sieben Ziffern benutzt und den Teilnehmern wurden jeweils sechs Additionsaufgaben gezeigt. Dieser Ansatz erwies sich allerdings als nicht aussagekräftig, weil nahezu alle Testpersonen alle Ziffernfolgen

richtig notiert hatten. In der Folge wurde dann auf sechs, sieben, acht und neun Ziffern erhöht. Die Anzahl der Additionsaufgaben wurde ebenfalls auf sieben angehoben.

Im Hauptteil des Testheftes, in dem sechs Aufgaben zum Kompetenzbereich Bewerten gestellt wurden, wurden die Aufgaben je nach teilnehmendem Kurs entweder auf Englisch oder auf Deutsch gestellt. Der Autor entschied sich bewusst gegen die Bereitstellung eines Glossars. Ein Ziel der Untersuchung ist es, eine Überforderung von Schülerinnen und Schülern bei der Bearbeitung komplexer Inhalte im bilingualen Biologieunterricht nachzuweisen. Diese Überforderung kann entweder inhaltliche oder sprachliche Gründe haben. Während die Antworten eine fachliche Überforderung abbilden, ist dies für eine sprachliche Überforderung nicht der Fall. Wenn ein Teilnehmer auf eine Teilaufgabe keine oder eine Nonsens-Antwort gibt, ist für den Autor nicht nachvollziehbar, ob diese dies aufgrund inhaltlicher oder sprachlicher Überforderung geschah. Beim Bereitstellen eines Glossars hingegen kann nicht überprüft werden, inwieweit die Teilnehmerinnen und Teilnehmer auf dieses zurückgreifen. Somit wird die Abbildung einer sprachlichen Überforderung erneut unmöglich. Wenn die Teilnehmer allerdings während der Testdurchführung nach Vokabelhilfe fragen müssen, ist es möglich diese zu gewähren, aber gleichzeitig eine sprachliche Überforderung zu notieren.

Die verwendeten Aufgaben wurden aus wissenschaftlichen Publikation zum Kompetenzbereich Bewerten entnommen und so angepasst und verändert, dass alle Teilkompetenzen von Bewertungskompetenz des Oldenburger Modells der Bewertungskompetenz (vgl. MITTELSTEN SCHEID/HÖBLE 2007; REITSCHERT ET AL. 2007) abgedeckt wurden (vgl. Abb.2). Aufgabe 1 behandelt das Dilemma einer ungewollten Schwangerschaft, die sich nicht mit den Lebensplänen der Beteiligten vereinbaren lässt. Dilemmata zeichnen sich gerade dadurch aus, dass es keine eindeutige Lösung gibt und dass sich mindestens zwei grundlegende Werte, die die Situation betreffen, so unvereinbar gegenüberstehen, dass es keine richtige oder falsche Lösung des Problems gibt (vgl. EGGERT & HÖBLE 2006). Für die Aufgabe im Testheft wurde auf die thematische Grundlage von DORIS ELSTER (vgl. 2011) zurückgegriffen. Die Bearbeitung der Aufgabe orientiert sich an der Sechs-Schritte-Methode von CORINNA HÖBLE, ist aber auf fünf Schritte reduziert. Der ursprüngliche vierte Schritt, die Unterscheidung zwischen deontologischen und utilitaristischen

Argumenten (vgl. HÖBLE/BÖGEHOLZ 2006), wurde nicht in die Aufgabe übernommen, weil für die Beantwortung ein gewisser Kenntnisstand von ethischem Basiswissen erforderlich ist, der nicht unbedingt bei den Schülerinnen und Schülern vorhanden ist. Diese Aufgabe deckt fünf Teilkompetenzen von Bewertungskompetenz ab, nämlich das Wahrnehmen und Bewusstmachen moralischer Relevanz, Ethisches Basiswissen, Urteilen, Argumentieren und Folgenreflexion. In der zweiten Aufgabe werden die Schüler aufgefordert, einen Syllogismus aufzustellen, indem sie eine deskriptive Prämisse mit einer moralisch-ethischen Prämisse verbinden und zu einem normativen Schluss kommen. Diese Aufgabe stammt ebenso wie die Aufgabe 3 von ELKE VISSER und CORINNA HÖBLE (2010). In Aufgabe 3 geht es erneut um das Wahrnehmen und Bewusstmachen moralischer Relevanz in verschiedenen Fragestellungen. Die Fragen selber müssen nicht beantwortet werden, die Teilnehmer müssen nur entscheiden, ob bei einer Beantwortung eine moralische Dimension vorliegt.

Aufgabe 4 lässt sich der Teilkompetenz Perspektivwechsel zuordnen. Die Teilnehmer müssen für zwei verschiedene Kunden den jeweils richtigen Apfel auswählen und begründen diese Auswahl exemplarisch an einem der Kunden. Diese Aufgabe ist nach der Aufgabe „Kaufentscheidung Apfelsorte“ der bik-Arbeitsgruppe Niedersachsen/Göttingen verändert worden (vgl. LEIBNIZ-INSTITUT FÜR DIE PÄDAGOGIK DER NATURWISSENSCHAFTEN 2009).

Aufgabe 5 widmet sich wieder einem bioethischen Thema. Das Krankheitsbild der Cystischen Fibrose wird vorgestellt und auf eine Heilungschance durch die Nutzung von embryonalen Stammzellen wird hingewiesen. In einer Teilaufgabe müssen die Teilnehmer beurteilen, welche Konsequenzen eine Legalisierung von Stammzellforschung hätte, was der Teilkompetenz Beurteilen entspricht. In einem weiteren Aufgabenteil sind die Teilnehmer mit einem Szenario konfrontiert, in dem zwei homozygote Träger des Gens für Cystische Fibrose aufgrund einer möglichen Erkrankung ihres Kindes über eine Abtreibung nachdenken. Hierbei müssen die Teilnehmer letztendlich entscheiden, ob das Paar eine Abtreibung vornehmen lassen sollte. Basierend auf ZOHAR & NEMET (2002) deckt diese Teilaufgabe die Teilkompetenzen Wahrnehmen und Bewusstmachen der eigenen Einstellung und Perspektivwechsel ab.

Die sechste Aufgabe wurde selbstständig konzipiert, basiert jedoch auf einem Wertepool, der im Zusammenhang mit dem Universitätsseminar „Bewerten im Biologieunterricht“ an der Universität Kassel im Wintersemester 2009/2010 erarbeitet wurde.

Für die Verwendung des Testheftes in den Kursen wurden vier verschiedene Versionen angefertigt, bei denen jeweils die Reihenfolge der Testaufgaben verändert wurde. In jedem Kurs wurden etwa gleich viele verschiedene Testheftversionen ausgeteilt. Dadurch konnte einerseits ein Abschreibeffekt verhindert werden, da die Teilnehmerinnen und Teilnehmer zu einem Zeitpunkt in der Regel mit einer anderen Testaufgabe beschäftigt waren. Andererseits sollten so zwei weitere Effekte von empirischen Untersuchungen verhindert werden. Bei Zeitmangel ohne Variation des Testheftes würde immer dieselbe Aufgabe unbeantwortet bleiben, was eine Auswertung dieser Anteile unmöglich macht. Andererseits sind bei längeren Tests Ermüdungseffekte erkennbar. Das bedeutet, dass die Aufgaben, die an hinterer Stelle im Testheft lokalisiert sind, in der Regel weniger erfolgreich bearbeitet werden. Durch die Variation der Aufgabenreihenfolge kann ausgeschlossen werden, dass Aufgaben aufgrund dieses Ermüdungseffekts unzureichend oder fehlerhaft beantwortet wurden. Die Variation stellt somit eine Abbildung der inhaltlichen oder sprachlichen Überforderung der Teilnehmer dar. Im Anhang ist Testheftversion A abgedruckt (Anhang 1).

4.3 Die Durchführung der Datenerhebung

Die Testhefte wurden persönlich vom Autor dieser Arbeit an die teilnehmenden Schülerinnen und Schüler ausgeteilt und das Projekt sowie die Durchführung wurde den Teilnehmern mündlich vorgestellt. In bilingual unterrichteten Kursen wurde das englischsprachige Testheft verteilt, die Teilnehmer, die Biologie in der Muttersprache belegten, erhielten entsprechend das Testheft in deutscher Sprache. Wie oben bereits erläutert wurden verschiedene Testheftversionen verteilt. Nach Klärung offener Fragen wurde der erste Teil des Testes durchgeführt. Der Autor führte sowohl die Zahlenfolgen, als auch die einfachen Additionsaufgaben in den jeweiligen Klassen vor. Besonderes Augenmerk lag darauf, dass die Teilnehmerinnen und Teilnehmer die Zahlenfolge wirklich erst nach der Ablenkungsphase notierten. Danach hatte die Schülerinnen und Schüler dann

insgesamt 60 Minuten Zeit, die sechs Aufgaben zu lösen und den Fragebogen am Ende auszufüllen. Während der Bearbeitung wurde darauf geachtet, dass die Teilnehmerinnen und Teilnehmer die Aufgaben selbstständig und ohne Abschreiben bearbeiteten. Der Autor der Arbeit stand während der gesamten Zeit für Fragen, insbesondere für Verständnis- und Vokabelfragen, zur Verfügung. Da bei der Konzeption des Testhefts auf ein Glossar bewusst verzichtet wurde, konnten sprachliche Probleme bei einzelnen Teilnehmern erkannt und auf dem Testheft notiert werden. Bei einem Glossar, deren Benutzung nicht überprüft werden kann, wäre diese präzise Beobachtung nicht möglich gewesen. Um sicherzustellen, dass jeder Teilnehmer das Testheft trotzdem möglichst vollständig bearbeiten konnte, wurden Vokabelfragen vor Ort beantwortet.

Nach der Bearbeitung der Testhefte wurden diese direkt wieder eingesammelt. Den Schülerinnen und Schülern, sowie dem Lehrer bzw. der Lehrerin wurde gedankt, offene Fragen wurden beantwortet. Nach Einsammeln der Arbeiten wurde auf jedem Testheft die Schule und die Lehrkraft notiert, um Besonderheiten, wie Schul- oder Lehrereffekte erkennen und beschreiben zu können.

In der Regel dauerte die Befragung der bilingualen Schülerinnen und Schüler deutlich länger als in muttersprachlichen Kursen. Das ist wahrscheinlich einer verlangsamten Lesegeschwindigkeit geschuldet, die vor allem bei den Aufgaben 1, 5 und 6 zum Tragen kommt.

Da mehrmals auf die Möglichkeit des Nachfragens hingewiesen wurde, kann davon ausgegangen werden, dass einzelne, nicht beantwortete Aufgaben das kognitive Leistungsvermögen der Teilnehmer überfordert hat. Wurden weite Teile des Fragebogens nicht oder nur mit Nonsense-Antworten beantwortet, lässt sich nicht ermitteln, ob diese Fehler Folge von fehlender Motivation, Desinteresse, Indifferenz oder eben Überforderung waren.

4.4 Die Auswertung

Die Auswertung erfolgte durch den Autor. Es wurde ein vorher erarbeitetes Codierungsschema verwendet, dem ein Erwartungshorizont zu Grunde liegt. Für die Aufgaben 1 und 5 wurden neben dem Erwartungshorizont auch die Niveaustufen der Bewertungskompetenz (vgl. MITTELSTEN SCHEID & HÖBLE 2007; REITSCHERT 2007)

herangezogen, während es für die Aufgaben 2, 3 und 4 klare Verhältnisse von richtigen und falschen Antworten gibt. Bei Aufgabe 6 musste bei einem Teil auch auf die Niveaus zurückgegriffen werden. Das Codierungsschema befindet sich im Anhang (Anhang 2). Mithilfe des Programmes Microsoft® Excel® wurde ein tabellarischer Datensatz erstellt, aus dem die Diagramme und Tabellen der Ergebnisdiskussion generiert wurden.

Während der Auswertung wurde beschlossen, die Aufgabe 6 des Testheftes nicht mit einzubeziehen. Die Aufgabe wurde zu unpräzise formuliert, sodass die Schülerinnen und Schüler eine zu große Bandbreite von Antworten gegeben haben. Während bei Aufgabe 3 der Weg der Bewertung mit FULL CREDIT, HALF CREDIT und NO CREDIT gewählt wurde, ist diese Methode hier nicht anwendbar. Würden die Schülerinnen und Schüler die Chance gehabt haben, ihre komplette Auswahl der relevanten Werte zu begründen, hätte Aufgabe 6 mit verschiedenen Niveaus bewertet werden können. Gefragt war aber nur eine Begründung für zwei gewählte Werte. Die Wahrscheinlichkeit, dass die Schülerinnen und Schüler für diese Teilaufgabe gerade jene Werte gewählt haben, die nicht eindeutig relevant sind, ist jedoch sehr gering.

Gleichzeitig bleibt anzumerken, dass diese Streichung vorgenommen werden kann, ohne dass eine Teilkompetenz von Bewertungskompetenz ungeprüft bleibt. Die betroffenen Teilkompetenzen „Argumentieren“ und „Wahrnehmen und Bewusstmachen moralischer Relevanz“ werden noch von andern Aufgaben überprüft.

5 Ergebnisse

5.1 Allgemeine Ergebnisse

Zu Beginn der Ergebnispräsentation sollen allgemeine Ergebnisse vorgestellt werden, die bei der Auswertung der Fragebögen ermittelt wurden und die für die Studie wichtig sind, ohne dass sie direkt eine der Fragestellungen beantwortet.

5.1.1 Anmerkungen zum Kognitionstest und Ergebnisse

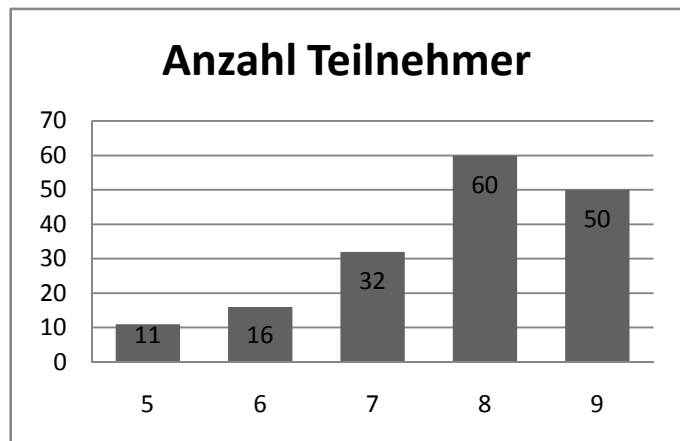
Beim Kognitionstest tauchte bei der Auswertung ein Problem auf. Der Test ist so angelegt, dass die Teilnehmer in vier Durchgängen eine willkürliche Zahlenfolge präsentiert bekommen, sich diese merken sollen, um sie nach einer Ablenkung

aufzuschreiben. In jedem Durchgang besteht die Zahlenfolge aus einer zusätzlichen Zahl, sodass die Zahlenfolge zuerst aus sechs, dann aus sieben, acht und schließlich neun Ziffern bestand. Natürlich änderte sich die Ziffernfolge jedesmal komplett und wurde nicht nur ergänzt. Ziel des Tests ist die Feststellung der Schwelle der kognitiven Überforderung. Nach MILLER (1956), müsste diese Schwelle bei ungefähr sieben *Chunks* liegen, also bei der zweiten Zahlenfolge. Bei der Auswertung fiel allerdings auf, dass es einige Schülerinnen und Schüler gab, die sich beispielsweise die erste, kürzeste Ziffernfolge nicht merken konnten, dann aber wiederum die letzten drei oder zwei korrekt wiedergeben konnten. Dieser Umstand erschwert die Festlegung einer Überforderungsschwelle. Der Test ist zusätzlich immer nur eine Momentaufnahme, die durch zahlreiche Faktoren, wie z.B. interne oder externe Ablenkung, Unaufmerksamkeit, Verpassen einer Ziffer durch eine physiologische Reaktion, wie Niesen oder Husten, etc., beeinflusst werden kann. Des Weiteren ist es möglich, dass einige Teilnehmer während des ersten Durchgangs das Verfahren nicht verstanden haben. Wenn ein Teilnehmer also beim ersten Durchgang die Zahlenfolge nicht richtig notiert hat, die nächste Zahlenfolge jedoch richtig notierte, wird davon ausgegangen, dass beim ersten Durchgang eine der oben angesprochenen Beeinträchtigungen vorliegt.

Schwieriger wird die Schwellenzuweisung bei Teilnehmern, die korrekte und nicht korrekte Zahlenfolgen im Wechsel notierten. Dies ist bei 45 Teilnehmern der Fall. Zwar wurden die Zahlenfolgen von einem internetbasierten Zufallsgenerator zusammengestellt, allerdings weisen die Zahlenfolgen der Durchgänge drei und vier eine gewissen Regelmäßigkeit auf, die sich die Teilnehmer gegebenenfalls besser merken konnten als die Zahlenfolgen der ersten beiden Durchgänge. Hier könnte eine Verfälschung vorliegen. Es wurde daher bestimmt, dass ein Teilnehmer, der eine Zahlenfolge falsch notiert, eine oder mehrere spätere jedoch wieder korrekt wiedergegeben hatte nicht die Überforderungsschwelle der letzten richtigen Kombination zugewiesen bekommt, sondern die der nächstrichtigen Zahlenfolge nach der falschen. Notiert ein Teilnehmer also richtig (6) – falsch (7) – richtig (8) – richtig (9), so erhält er die Schwelle 8, weil es die nächste richtige Notierung nach dem Fehler war. Er kann deswegen nicht die Schwelle 9 bekommen, weil er anders als ein Teilnehmer der die Schwelle 9 zugewiesen bekommt eben nicht alle Zahlenfolgen korrekt notiert hat. Diese Lösung ist zwar nicht optimal, sie ist

allerdings dem Aufbau des Tests geschuldet und bezieht auch den Aspekt mit ein, dass jede notierte Zahlenfolge nur eine Momentaufnahme ist. Nach der Anwendung dieser Methode bleiben noch 13 Teilnehmer ohne Schwelle über. Bei diesen Sonderfällen handelt es sich um Teilnehmer, die entweder nur eine Zahlenfolge richtig beantwortet haben, oder die der Folge richtig (6) – falsch (7) – falsch (8) – richtig (9) folgen. Im ersten Fall wird bei der richtigen Beantwortung von einem Zufall ausgegangen, weshalb die Schwelle 5 zugeteilt wird. Bei dem anderen Phänomen wird die Stufe 7 zugeteilt, da der Teilnehmer diese Zahlenfolge zwar nicht richtig notierte, dafür aber eine längere, die sich jedoch durch eine gewisse Wiederholung der Ziffern auszeichnete und damit vermeintlich einfacher war.

Jeder Teilnehmer bekommt nach Auswertung seines jeweiligen Kognitionstest eine Überforderungsschwelle zugeordnet, die die Anzahl der *Chunks* widerspiegelt, die er oder sie verarbeiten kann, ohne überfordert zu werden. Analog zu der Länge der Zahlenfolge wurden die Schwellen 6,7,8 und 9 zugeordnet. Wurde von einem Teilnehmer keine Zahlenfolge richtig notiert, wird eine Überforderungsschwelle von fünf *Chunks* vermutet, weshalb die Schwelle 5 verwendet wird. Durchschnittlich konnten sich die Teilnehmer 7,73 Ziffern lange Zahlenfolgen merken.



Damit liegt das Ergebnis etwas oberhalb der von Miller (1956) **Abbildung 8: Anzahl der Teilnehmer pro ermittelte Überlastungsschwelle (n=169)**

ermittelten Werte, ohne jedoch deutlich davon abzuweichen. Die Abweichung ist sicherlich den oben erwähnten artifiziellen Anpassungen geschuldet. Insgesamt haben 11 Teilnehmer eine Überforderungsschwelle von fünf *Chunks*, 16 Teilnehmer haben die Schwelle bei sechs *Chunks*, 32 Teilnehmer können sich sieben *Chunks* merken, ohne kognitiv überfordert zu sein. Auf die Schwelle bei acht bzw. neun *Chunks* entfallen 60 bzw. 50 Teilnehmer. Der Kognitionstest liefert also ein heterogenes Ergebnis, das gut mit den ermittelten Bewertungskompetenzen verglichen werden kann.

5.1.2 Ergebnisse der Testheftaufgaben

Bevor für die Auswertung die Korrelation der kognitiven Leistungsfähigkeit der Teilnehmer und ihrer Bewertungskompetenz zu untersuchen, sollen die Aufgaben des Testheft zuerst hinsichtlich ihrer Ergebnisse bei der Befragung ausgewertet werden. In Aufgabe 1 geht es um ein Dilemma mit ungewollter Teenager-Schwangerschaft, die nicht mit den Lebensträumen der potentiellen Eltern vereinbar ist. Desweiteren spielen die Religiosität der Eltern des Mädchens und fehlende finanzielle Ressourcen eine Rolle. Die Lösung der ersten Aufgabe orientiert sich wie oben beschrieben an der Sechs-Schritte-Methode von CORINNA HÖBLE (HÖBLE & BÖGEHOLZ 2006). In Aufgabenteil A sollten die Schülerinnen und Schüler das Dilemma beschreiben. Hierbei wurden folgende Niveaustufen definiert:

Tabelle 2: Definition der Niveaustufen zu den Antworten zu den Aufgabe 1A und 5D

Niveaustufe	Inhalt
0	keine, oder falsche Beschreibung
I	Deskriptiv-pragmatische Problemwahrnehmung
II	Deskriptive Problemwahrnehmung mit Bewusstmachung des ethisch-moralischem Problems, möglicherweise emotionale
III	Objektive Erkenntnis des moralisch-ethischen Wertkonflikts

Bei der Bearbeitung der Aufgabe 1A hat der Großteil der Teilnehmer (139) nur Niveaustufe I erreicht, und damit eine deskriptiv-pragmatische Problemwahrnehmung gezeigt. Typische Antworten enthielten die Darstellung der ungewollten Schwangerschaft, sowie der unklaren finanziellen Situation. Ein ethisches Dilemma zwischen den Werten „Leben des ungeborenen Kindes“ und „Glück/Selbstentfaltung der Eltern“ wurde nur in den seltensten Fällen (6) getroffen. Etwas häufiger (22) wurde eine Antwort mit emotionaler Färbung gegeben und zwei Teilnehmer gaben

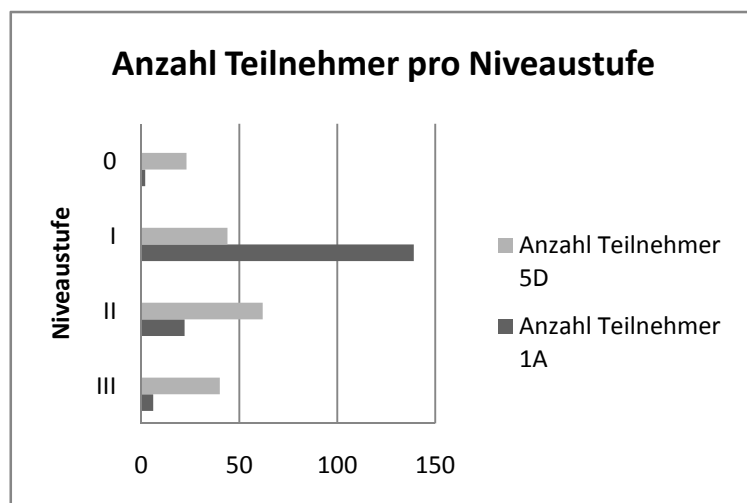


Abbildung 9: Anzahl der Antworten der Teilnehmer pro Niveaustufe für die Aufgabenteile 1A und 5D (n=169)

eine Antwort mit falscher oder nicht vorhandener Beschreibung (vgl. Abb. 9). Die gleichen Niveauabstufungen wurden auch für die Aufgabe 5D zu Grunde gelegt. Hierbei besteht das Dilemma darin, dass zwei heterozygote Träger der Erbinformation für zystische Fibrose ein Kind bekommen könnten und nun über eine Abtreibung nachdenken. Hierbei gaben zwar mehr Teilnehmer (23) eine falsche Antwort, dafür war aber auch die Zahl der Schülerinnen und Schüler, die die Niveaustufen II und III erreichten ebenfalls mit je 62 bzw. 40 auch deutlich höher, während 44 Teilnehmer eine Antwort notierten, die der Niveaustufe I zuzuordnen ist. (vgl. Abb. 9). Es ist zu vermuten, dass das Dilemma im Fall der Aufgabe 5D einfacher zu identifizieren war, als in der ersten Aufgabe. Das Dilemma besteht in diesem Fall zwischen den Werten „Leben des ungeborenen Kindes“ und „Gesundheit/Ersparung von Leid“.

Bei den Aufgaben 1B und 1C sollten Lösungen für das Dilemma bzw. Werte genannt werden, die bei dem Dilemma eine Rolle spielen. Die Niveaustufen beziehen sich in beiden Teilaufgaben auf die Anzahl der Optionen bzw. Gründe die genannt wurden. Bei der Aufgabe 1B nannten 15 Teilnehmer keine Option, die das Dilemma lösen könnte oder nannten Optionen, die das Dilemma erst gar nicht entstehen lassen hätten, wie „Verhütung“ (Niveaustufe 0). 100 Schülerinnen und Schüler nannten zwei Optionen, in den meisten Fällen „Geburt und Abtreibung“ (Niveaustufe I). Wenn zudem noch weitere Optionen wie z.B. „Freigabe zur Adoption“ oder „Kinderklappe“ genannt wurden, so entsprach dies der Niveaustufe II, was bei 54 Teilnehmern der Fall war. Bei der Aufgabe 1C nannten 30 Teilnehmer keinen relevanten Wert (Niveau 0), 64 Teilnehmer nannten einen relevanten Wert (Niveau I), während 56 Teilnehmer zwei relevante Werte (Niveau II) und 19 Teilnehmer drei oder mehr relevante Werte angaben (Niveau III). Bei der Auswertung wurden auch solche Antworten als Wert gezählt, die den Wert nicht explizit nannten, sondern diesen umschrieben. Diese Entscheidung wurde getroffen, weil während der Durchführung der Untersuchung vermehrt die Frage aufkam, was ein ethischer Wert sei.

Bei der Aufgabe 1D ging es darum eine Entscheidung zu treffen hinsichtlich der Frage, wie sich das Teenager-Paar aus dem Dilemma verhalten soll. Als Auswahlmöglichkeiten dienten dabei die im Aufgabenteil 1B genannten Optionen. Auch hierbei wurde wiederum eine Niveauabstufung vorgenommen (vgl. Tabelle 3).

Tabelle 3: Definition der Niveaustufen zu den Antworten zu der Aufgabe 1D

Niveaustufe	Inhalt
0	es findet keine Aussage statt.
I	Der Schüler macht eine zusammenhangslose Aussage, die zudem keine Begründung enthält.
II	Der Schüler bezieht klar Stellung, begründet seine Meinung aber nicht oder kaum.
III	Der Schüler bezieht klar Stellung und begründet seine Meinung mit einem logisch konsistenten Argument

Auf die Niveaustufe 0 entfielen fünf Teilnehmerantworten, während 25 Schülerinnen und Schüler eine zusammenhangslose Aussage machten oder keine Entscheidung trafen. Unter dieser Niveaustufe wurden auch Antworten gewertet, die die verschiedenen Optionen gegeneinander abwägen, weil eine definitive Entscheidung bei dieser Aufgabe im Vordergrund steht. Auch bei der Sechs-Schritt-Methode von HÖBLE (HÖBLE & BÖGEHOLZ 2006) wird dies so formuliert. Aussagen wie „Die Entscheidung kann ich nicht fällen und ich hoffe, dass ich

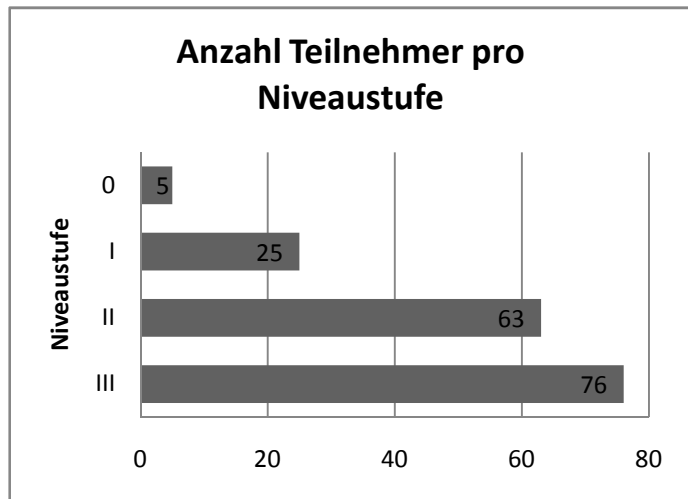


Abbildung 10: Anzahl der Antworten der Teilnehmer pro Niveaustufe für die Aufgabenteile 1D (n=169)

niemals in eine solche Situation komme“ lassen zwar erkennen, dass der Teilnehmer eine gewisse Kompetenz beim Erkennen und Bewusstmachen von ethischen Dimensionen aufweisen, allerdings verfügen sie nicht über die Teilkompetenz „Urteilen“. Eine Auswertung der Ergebnisse bezüglich eines Vorhandenseins der Teilkompetenzen des Oldenburger Modells (vgl. REITSCHERT et al. 2007; MITTELSTEN-SCHEID & HÖBLE 2007) erfolgt in Abschnitt 5.1.3. Gleichzeitig wiesen 63 Antworten die Niveaustufe II und 76 Antworten die Niveaustufe III auf. Der Unterschied zwischen diesen beiden Niveaustufen besteht vorrangig in der argumentativen Untermauerung des Urteils. Hierbei kommt es auf die Anzahl der

Argumente, deren Bedeutung und Sinnhaftigkeit an; besonders ausschlaggebend für Niveau III ist die Integration von Werten in der Entscheidungsfindung.

Die erste Aufgabe wird abgeschlossen durch eine Folgenreflexion. Die Teilnehmer sollten die Folgen ihrer Entscheidung für das Teenager-Paar darstellen. Hierbei wurden wiederum drei Kompetenzniveaus unterschieden. Bei Niveau 0 wurden keine Folgen genannt oder keine sinnvollen bzw. wahrscheinliche Folgen. Niveau I und II unterscheiden sich dadurch, ob Primär-, oder auch Sekundärfolgen für unmittelbar und mittelbar betroffene Personen aufgelistet wurden. Unmittelbare Personen sind in dem Kontext Linda, Bob und das ungeborene Kind, mittelbare Personen sind die Eltern des Paares. Primärfolgen sind Tod oder Leben des Babys, Mutterschaft und Schulabbruch, während die Sekundärfolgen sich auch auf Langzeitfolgen wie Familiengründung, Schuldgefühle und zerrüttetes Verhältnis zu den Eltern erstrecken. Folglich erreichten 10 Schülerinnen und Schüler Niveau 0, 105 Teilnehmer erreichten Niveau I; 54 Teilnehmer erreichten Niveau II.

Bei den Aufgaben 2 bis 4, mit Ausnahme des zweiten Teils der Aufgabe 4, gibt es keine Niveauabstufungen, sondern nur FULL CREDIT oder NO CREDIT mit Ausnahme der Aufgabe 3, bei der auch HALF CREDIT möglich ist. Dies erklärt sich dadurch, dass die Schülerinnen und Schüler bei Aufgabe 3 in zehn Unterrichtssituationen entscheiden müssen, ob diese eine ethisch-moralische Dimension haben, oder ob bei der Vermittlung vorrangig Fachwissen vermittelt werden muss. FULL CREDIT bekommt man nur, wenn man alle zehn Situationen richtig einschätzt (s. Anhang 3). Da aber ein Schüler, der zwei Fehler macht eine höhere Teilkompetenz „Wahrnehmen und Bewusstmachen ethischer Relevanz“ aufweist, als ein Schüler, der 8 Fehler hat, wurde die Verwendung der Bewertungseinheit HALF CREDIT beschlossen. Teilnehmer, die bis zu drei Fehler haben, fallen in diese Kategorie. Nach diesem Bewertungsschema erhalten 54 Teilnehmer NO CREDIT, 4 FULL CREDIT und 111 HALF CREDIT.

In Aufgabe 2 mussten die Schülerinnen und Schüler aus vorgegebenen Statements insgesamt zwei Syllogismen aufstellen, bei denen aus einer ethischen und einer deskriptiven Prämisse ein normativer Schluss, bzw. eine Handlungsaufforderung hervorgeht. Aufgrund der Komplexität werden die beiden Syllogismen einzeln gewertet. Dabei stellten 80 Teilnehmer beide Syllogismen falsch und 48

Schülerinnen und Schüler zumindest einen Syllogismus richtig auf, während 41 Teilnehmer beide Syllogismen richtig aufstellen konnten.

Beim ersten Teil der Aufgabe 4 mussten die Teilnehmer auf der Grundlage einiger Informationen über vier verschiedene Apfelsorten, die jeweils richtige Apfelsorte für einen Kunden auswählen, der ein bestimmtes Anforderungsprofil hatte. Hierbei bekommt man als Teilnehmer nur FULL CREDIT, wenn man für beide Kunden den richtigen Apfel nennt. Mehrfachnennungen von Apfelsorten bei einem oder beiden Kunden führt ebenso zu NO CREDIT, wie eine Falschnennung. Diese Bewertung resultierte in 125 Antworten, die FULL CREDIT bekamen, gegenüber 44-mal NO CREDIT. Die Ergebnisse der Aufgaben 2A, 2B, 3 und 4A sind in Abbildung 11 zusammengefasst. Die Aufgaben 2A und 2B sind darin zusammengefasst und die Wertung HALF CREDIT entspricht einem richtigen und einem falschen Syllogismus.

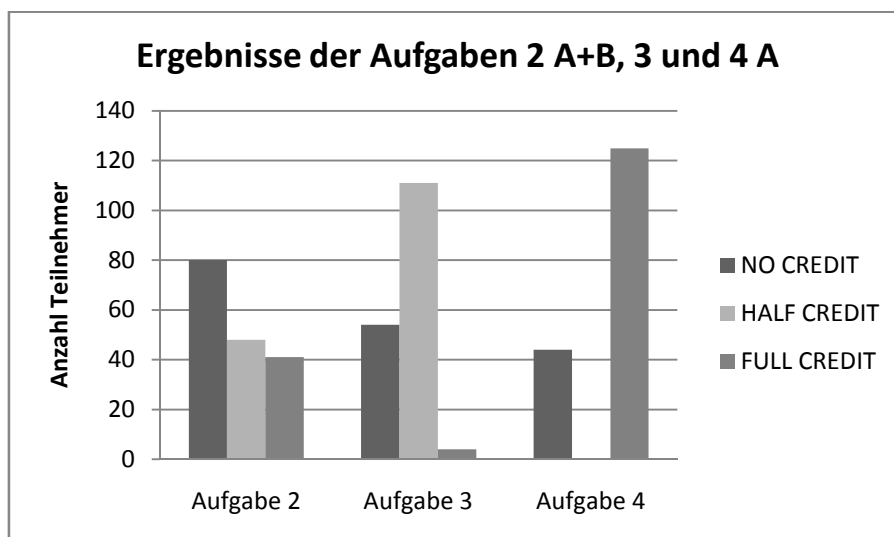


Abbildung 11: Ergebnisse der Aufgaben 2A, 2B, 3 und 4A (n=169)

Bei Aufgabe 4B hingegen gibt es wieder Niveaustufen der Bewertungskompetenz, ähnlich wie bei den Aufgaben 1B und 1C definieren sich die Niveaustufen über die Anzahl der Argumente für den Erwerb von Bio-Äpfeln aus lokalem Anbau. Niveau 0 (6) entspricht keiner Nennung von Gründen oder eine Nennung von Nonsens-Gründen, Niveau I (35) entspricht einem Grund, Niveau II (86) entspricht 2 Gründen und Niveau 3 (42) entspricht drei und mehr Gründen für den Kauf. Die Zahlen in Klammern bezeichnet jeweils die Anzahl der Schülerinnen und Schüler, deren Antwort dem jeweiligen Niveau entsprechen.

Abschließend gilt es, die Ergebnisse der Aufgabe 5 zu betrachten. Bei den Teilaufgaben 5A, 5C und 5E definieren sich die Niveaustufen wieder über die Anzahl der genannten Gründe, Argumente oder Werte. Niveau 0 entspricht dabei immer keiner Nennung oder einer Nonsense-Nennung. Niveau I wird über die Nennung von einem Grund, einem Argument oder einem Wert definiert. Für Niveau II müssen dann zwei Nennungen erfolgen und für Niveau III dann analog dazu mehr als zwei. Bei Teilaufgabe 5C gibt es kein Niveau III. Die Ergebnisse der Aufgaben sind in Tabelle 4 dargestellt.

Tabelle 4: Ergebnisse der Teilaufgaben 5A, 5C und 5E

Niveau	Teilaufgabe 5A	Teilaufgabe 5C	Teilaufgabe 5E
0	39	19	24
I	60	74	55
II	57	76	78
III	13	---	12

In Aufgabe 5B werden die Schülerinnen und Schüler aufgefordert, mögliche Konsequenzen einer Legalisierung von Stammzellenforschung zu nennen. Der niveaudefinierende Aspekt ist hierbei der Abstraktionsgrad der Konsequenzen. Wenn ein Schüler oder eine Schülerin eine Konsequenz nennt, die als persönlich beschrieben werden kann, so entspricht diese Antwort Niveau I. Sie ist deswegen geringer zu bewerten, weil sie einerseits aus dem Aufgabentext ableitbar ist und andererseits auf einer sehr persönlichen Basis ansetzt. Eine typische Antwort wäre: „Durch die Stammzellenforschung kann das Leiden von Kranken gelindert werden“. Auf Niveau II hingegen werden weitreichendere, gesellschaftlich relevante Konsequenzen genannt, wie beispielsweise: „Durch Stammzellenforschung kann die Medizin große Fortschritte machen, allerdings besteht die Gefahr des Missbrauchs.“ Am höchsten ist eine Konsequenzen-Reflektion zu bewerten, bei der persönliche und gesellschaftliche Konsequenzen genannt werden (Niveau III). Jegliche Nicht-Nennung oder Nonsense-Nennung wird mit Niveau 0 gewertet.

Die Teilaufgabe 5D wurde schon weiter oben im Vergleich mit der Aufgabe 1A präsentiert. Die Niveaustufen der Teilaufgabe 5F entsprechen denen der Aufgabe 1D (vgl. Tab. 3). Es muss auch hier wieder eine Entscheidung in einer Dilemma-Situation getroffen werden, wobei Antworten ohne klare Entscheidung als Niveau I gewertet wurden, was bei 43 Schülerinnen und Schülern der Fall war, während

Niveau 0 einer Antwort entspricht, die entweder gar nicht gegeben wurde, oder nicht zum Thema passte. Dieses Niveau entfiel auf 29 Teilnehmer. Für Niveau II und III mussten die Antworten eine klare Positionierung der Schülerin oder des Schülers enthalten. Bei Niveau II wird diese aber nicht oder kaum mit Argumenten belegt. Dieses Niveau wurde von 48 Teilnehmern erreicht, während 49 Schülerinnen und Schüler ihre Entscheidung ausreichend begründeten und damit Niveau III erreichten.

In der Abbildung 12 sind die Ergebnisse der Aufgabe 5, ohne die Teilaufgabe 5D abgebildet.

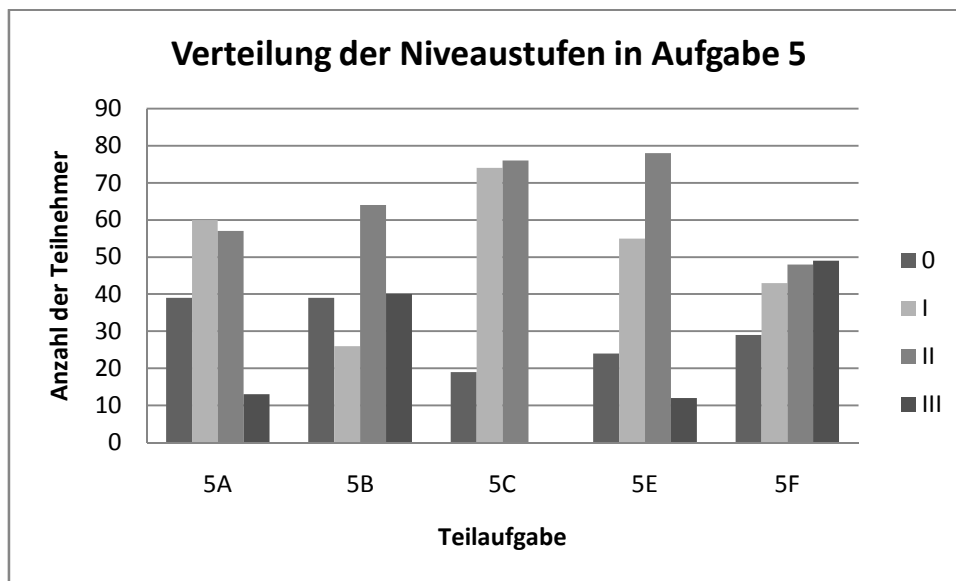


Abbildung 12: Ergebnisse der Aufgabe 5, außer Teilaufgabe D; Verteilung der Ergebnisse auf die Niveaustufen (n=169)

5.1.3 Ergebnisanalyse hinsichtlich der Teilkompetenzen nach dem Oldenburger Modell von REITSCHERT ET AL.

Mit Blick auf die Bewertungskompetenz erscheint es sinnvoll, die Ergebnisse der Aufgaben des Testheftes hinsichtlich der einzelnen Teilaspekte des Oldenburger Modells der Bewertungskompetenz von REITSCHERT ET AL. (2007) darzustellen. In Tabelle 5 werden die prozentualen Anteile der Schülerantworten pro Niveaustufe zusammengefasst. Über jeder Teilaufgabe ist die jeweilige Teilkompetenz, oder im Fall der Aufgaben 1D und 5F die Teilkompetenzen, vermerkt. Des Weiteren wurde der jeweils höchste Wert markiert, um eine Aussage über die durchschnittlich erreichten Niveaus zu treffen.

Auffällig ist, dass die meisten Antworten bei den Teilfragen in einem mittleren Niveaubereich liegen. Addiert man die Prozentzahlen der Niveaus I und II, bzw. betrachtet man den Prozentsatz der Antworten auf Niveau I bei solchen Teilfragen, bei denen es nur drei Niveaustufen gibt, so liegen Werte von 43,8% (Aufgabe 5C) bis 95,2% (Aufgabe 1A) vor. Bei den drei Teilaufgaben, die die Teilkompetenz „Argumentieren“ betreffen, nämlich 1D, 5C und 5F, liegt die Mehrheit der Schülerantworten auf dem höchsten Niveau. REITSCHERT ET AL. (2007) weisen darauf hin, dass die Teilkompetenz „Argumentieren“ und auch die Teilkompetenz „Perspektivwechsel“ nicht von den anderen Teilkompetenzen getrennt werden kann. Darüber hinaus ist besonders die Teilkompetenz „Argumentieren“ auch Bestandteil anderer Kompetenzbereiche, z.B. Kommunikation. Auch in anderen Schulfächern spielt das Argumentieren eine große Rolle, weshalb hier eine hohe Kompetenz zu erwarten war. Oftmals sind die Teilkompetenzen „Argumentieren“ und „Urteilen“ nicht voneinander zu trennen, weshalb sie auch bei den Testheftfragen 1D und 5F zusammen abgefragt werden. Bei beiden Teilaufgaben erreichte Mehrheit der Teilnehmer Niveau III (44,9% bzw. 29,0%). In Aufgabe 4A müssen die Schüler auf der Basis von Informationen ebenfalls ein Urteil fällen. Dieses Urteil beinhaltet jedoch kein ethisch-moralisches Moment, was die hohe Erfolgsquote der Schülerinnen und Schüler bei dieser Aufgabe erklärt. Immerhin wurden 74,0% der Schülerantworten mit FULL CREDIT gewertet, während 26,0% der Antworten falsch waren und somit mit NO CREDIT bewertet wurden.

Von besonderem Interesse ist auch der Vergleich der erreichten Niveaustufen bei zwei Teilaufgaben, die die gleiche Teilkompetenz der Bewertungskompetenz abdecken und abfragen. Ethisches Basiswissen ist sowohl in Aufgabe 1C als auch 5A von hoher Bedeutung. In beiden Teilaufgaben mussten die Schülerinnen und Schüler relevante Werte nennen, die einen Einfluss auf eine Situation haben. Schon bei der Auswertung fiel auf, dass nur die wenigsten Teilnehmer Werte im Sinne der Ethik formulierten. Dies, sowie die erreichten Kompetenzniveaus geben einen Hinweis darauf, dass zwar die meisten Schüler über ein gewisses ethisches Grundwissen verfügen, welches allerdings noch ausbaubar ist.

Tabelle 5: Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, nach Teilaufgaben; in %, höchster Wert markiert

	Wahrnehmen und Bewusstmachen moralischer Relevanz	Beurteilen	Ethisches Basiswissen	Urteilen / Argumentieren	Folgenreflexion	Perspektivwechsel	Ethisches Basiswissen	Beurteilen	Argumentieren	Wahrnehmen und Bewusstmachen moralischer Relevanz	Perspektivwechsel	Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren
Niv.	1A	1B	1C	1D	1E	4B	5A	5B	5C	5D	5E	5F
0	1,2	8,9	17,8	3,0	6,0	3,6	23,1	23,1	11,2	13,6	14,2	17,2
I	82,2	59,2	37,9	14,8	62,1	20,7	35,5	15,4	43,8	26,0	32,5	25,4
II	13,0	31,9	33,1	37,3	31,9	50,9	33,7	37,9	45,0	36,7	46,2	28,4
III	3,6	---	11,2	44,9	---	24,8	7,7	23,6	---	23,7	7,1	29,0

In beiden Teilaufgaben entfällt die Mehrheit der Schülerleistungen auf die Niveaus I und II, was einer Nennung von einem bzw. zwei relevanten ethischen Werten entspricht. Diese Werte sind allerdings nahezu direkt aus dem Text ableitbar.

Im Fall einer Abtreibung ist der Wert „Leben“ beispielsweise bereits impliziert. Nur jene 11,2% bzw. 7,7%, die in diesen Teilaufgaben das Niveau III erreichten haben ethisch relevante Werte genannt, die sich nicht auf den ersten Blick erschließen. Um das ethische Basiswissen der Schülerinnen und Schüler zu erhöhen, muss auch im Religions- oder Ethikunterricht auf biologische Dilemmata zurückgegriffen werden. Auf die Bedeutung dieser Studie für den Biologieunterricht wird jedoch in der Abschlussdiskussion eingegangen.

Tabelle 6: Anteil der Stichprobe an den Wertungseinheiten FULL CREDIT, HALF CREDIT und NO CREDIT nach Teilaufgaben; in %, höchster Wert markiert

	Ethisches Basiswissen	Wahrnehmen und Erkennen ethischer Relevanz	Beurteilen
Wertung	2	3	4A
FULL CREDIT	24,3	3,0	74,0
HALF CREDIT	28,4	65,1	---
NO CREDIT	47,3	31,9	26,0

Auch in der Aufgabe 2 wird das ethische Basiswissen abgefragt. Die Schülerinnen und Schüler müssen aus verschiedenen Aussagen einen Syllogismus bilden. Diese Technik geht auf die ethische Logik von Aristoteles zurück, wodurch bei dieser Aufgabe das ethische Basiswissen nicht nur inhaltlich, sondern auch methodisch abgefragt wird. Bei dieser Aufgabe schnitten die Schülerinnen und Schüler relativ schlecht ab (vgl. Tab. 6). Mit 47,3% haben knapp die Hälfte der Teilnehmer NO CREDIT bei der Aufgabe erreicht, was deswegen bedenklich ist, da die Aufgabe lediglich darin bestand, vorgegebene Statements zu kombinieren. Weitere 28,4% haben HALF CREDIT erreicht. Die Aufgabe 2 besteht aus zwei Syllogismen, das bedeutet, dass diese Gruppe immerhin einen Syllogismus richtig kombiniert hat. Ein Syllogismus kombiniert eine ethische und eine deskriptive Prämisse zu der normativen Aufforderung, Eizellspende in Deutschland zu legalisieren, der andere Syllogismus fordert die Beibehaltung des Verbots der Eizellspende in Deutschland. Weit über 90% der Teilnehmer, die HALF CREDIT erreichten, haben den positiven Syllogismus richtig kombiniert. Nur knapp ein Viertel (24,3%) der Teilnehmer hat die beiden Statements korrekt kombiniert.

Das ethische Basiswissen hat wiederum auch einen großen Einfluss auf die Fähigkeit, eine ethisch-moralische Relevanz wahrzunehmen und sich bewusst zu machen. Betrachtet man die Ergebnisse für diese Teilkompetenz, welche in Aufgabe 1A und 5D abgefragt wurden, fällt unmittelbar auf, dass sich die Ergebnisse stark unterscheiden (vgl. Tab. 5). Bei Aufgabe 1A wurde das Dilemma von 82,2% der Teilnehmer nur deskriptiv-pragmatisch beschrieben. Daraus kann gefolgert werden, dass diese Situation nicht als ethisches Dilemma wahrgenommen wird. Diese Einschätzung lässt sich auch bei der Nennung relevanter Werte beobachten: hier wurde hauptsächlich nur der Wert des „Lebens“ genannt. Werte, wie „Glück“, „Recht auf Entfaltung“ und „Freiheit“, die sich auf die werdenden Eltern beziehen, wurden nur selten genannt. Anders stellt sich die Situation bei Aufgabe 5D dar. Hier sind die Schülerantworten regelmäßiger auf die Niveaustufen verteilt. Die beschriebene Situation wurde öfter und stärker als ethisches Dilemma erkannt und beschrieben. Die Aufgabe 3 bietet einen weiteren Datensatz zu dieser Teilkompetenz (vgl. Tab. 6). Bei dieser Aufgabe haben nur 3% der Teilnehmer FULL CREDIT erreicht, was bedeutet, dass sie alle Situationen hinsichtlich ihrer ethisch-moralischen

Relevanz richtig bewertet haben. Die Mehrheit von 65,1% haben eine bis drei Situationen falsch bewertet, was der Wertung HALF CREDIT entspricht, während die restlichen 39,9% mehr als drei Situationen falsch bewertet haben, was in einer NO CREDIT-Wertung resultiert.

Die Teilkompetenz „Folgenreflexion“ wurde nur in einer Teilaufgabe direkt abgefragt, wobei mit 62,1% aller Schüler nur Primärfolgen für die unmittelbar und mittelbar betroffenen Personen nannten (vgl. Tab. 5). Wenn im Biologieunterricht allerdings eine Bewertungskompetenz erreicht werden soll, so müssen die Schülerinnen und Schüler auch in der Lage sein über Folgen zu reflektieren, die in einem breiteren Kontext oder zeitlich versetzt auftreten. Die große Bedeutung einer Folgenreflexion besteht gerade darin, dass die Bedeutung einer Entscheidung in einem kleinen, möglicherweise persönlichen, Rahmen für die Gesamtgesellschaft abgeschätzt und betrachtet wird. Auch hier besteht Nachholbedarf im schulischen Kontext. In Aufgabe 5B wurden die beiden Teilkompetenzen „Folgenreflexion“ und „Beurteilen“ abgefragt. Hierbei erreichten mehr als 50% der Teilnehmer das Niveau II oder höher.

Erfreulich sind die Ergebnisse für die Teilkompetenz „Perspektivwechsel“. In beiden Teilaufgaben, die diese Teilkompetenz abfragen, liegen mehr als 50% der Antworten auf Niveau II oder III. Ähnlich wie schon bei der Teilkompetenz „Argumentieren“, ist auf die Fähigkeit zum Perspektivwechsel für andere Teilkompetenzen unerlässlich und wird auch außerhalb des Biologieunterrichts regelmäßig eingeübt. Dennoch sollte das weitere Entwicklungspotential bei dieser Teilkompetenz nicht unterschätzt werden.

Bleibt noch die Teilkompetenz „Beurteilen“. Anders als bei der Kompetenz „Urteilen“ geht es hierbei nicht darum, eine begründete Entscheidung zu treffen, die durch Argumente gestützt wird. Das Beurteilen stellt allerdings eine wichtige Voraussetzung für die Entscheidungsfindung voraus. Die Schülerinnen und Schüler sind in den Aufgaben 1B und 5B aufgefordert eine Situation hinsichtlich einer Fragestellung zu beurteilen, die maßgeblich für die Entscheidungsfindung ist, weil mögliche Handlungsoptionen (Teilaufgabe 1B) und mögliche Konsequenzen (Teilaufgabe 5B) erörtert werden sollen. Die Ergebnisse für die Teilaufgabe 5B sind etwas besser als bei 1B, was aber auch daran liegt, dass bei dieser nur drei

Niveaustufen angewendet werden, während jene mit vier Niveaustufen weiter differenziert (vgl. Tab. 5).

5.2 Ergebnisse zu den Unterfragestellungen

Auf der Grundlage der vorher präsentierten Ergebnisse kann man letztendlich eine Zusammenfassung erstellen, die die Korrelation darstellt, zwischen der kognitiven Leistungsfähigkeit der Schülerinnen und Schüler und der Ausprägung der Bewertungskompetenz, die diese im Testheft an den Tag gelegt haben. Hierzu werden die Schwellenwerte der kognitiven Leistungsfähigkeit, die in Kapitel 5.1.1 präsentiert wurden mit einem Bewertungskompetenzwert, der in der Folge beschrieben wird, korreliert.

Der erwähnte Bewertungskompetenzwert ist eine Zahl, die auf den Antworten zu den Aufgaben im Testheft basiert. Jeder Schüler bekommt folglich eine Zahl zugewiesen. Je höher diese Zahl ist, desto höher ist auch die Bewertungskompetenz des jeweiligen Schülers einzuschätzen. Um den Wert zu generieren, werden den Niveaustufen verschieden hohe Werte zugewiesen. Die Einzelwerte aller Teilaufgaben eines Schülers werden addiert und ergeben schließlich den Wert. Dabei entsprechen die Niveaustufen auch den Punkten der Bewertungskompetenz. Erreicht ein Schüler bei einer Teilaufgabe die Niveaustufe III, so erhält er dafür drei Punkte. Bei den Aufgaben, bei denen es keine Niveaustufen, sondern nur die Ergebnisse FULL CREDIT und NO CREDIT gibt, erreicht man durch ersteres einen Punkt, während NO CREDIT zu keinem Punkt führt. In Aufgabe 3 gibt es des Weiteren die Bewertungsstufe HALF CREDIT, die in diesem Fall mit einem Punkt zu Buche schlägt, während FULL CREDIT dabei mit zwei Punkten gewertet wird. Der maximal erreichbare Kompetenzwert ergibt sich aus der Addition aller Punkte der Teilaufgaben und liegt bei 38.

Kombiniert man nun die Ergebnisse des Kognitionstest und die Ergebnisse der Aufgaben und die dabei erreichten Niveaustufen von Bewertungskompetenz, so lässt sich die Korrelation dieser beiden Aspekte ermitteln und darstellen. Für diese Arbeit wird der Korrelationskoeffizient von Pearson verwendet, der in einem Bereich von +1 bis -1 liegen kann, wobei positive Werte dem Paradigma „Je mehr X, desto mehr Y“ entsprechen, während negative Werte „Je mehr X, desto weniger Y“ bedeuten. Durch den ermittelten, einheitenlosen Wert kann der Zusammenhang zwischen zwei

Merkmale beschrieben werden. Folgende Tabelle gibt einen Überblick über die Aussage des Korrelationskoeffizienten r .

Tabelle 7: Aussagekraft des Korrelationskoeffizienten r (nach Pearson)

$r = 1$	$r = .99 - .8$	$r = .79 - .3$	$r = .3 - .01$	$r = 0$
Perfekte Korrelation	Hohe Korrelation	Mittlere Korrelation	Geringe Korrelation	Keine Korrelation

Für die Errechnung der Korrelation in dieser Arbeit wurde bei Microsoft® Office Excel® die Funktion =KORREL($Ax_1:Ax_n;Bx_1:Bx_n$) verwendet, mit der man beliebig viele Merkmalspaare korrelieren kann.

Bevor die Gesamtkorrelation der kognitiven Leistungsfähigkeit mit der Bewertungskompetenz, über die Werte Kognitionswert und Kompetenzwert, berechnet wird, sollen die Ergebnisse allerdings noch in Bezug auf die Unterfragestellungen dargestellt werden, bevor die Korrelation am Ende die übergeordnete Fragestellung dieser Arbeit beantwortet.

5.2.1 Unterscheidung nach Alter bzw. Jahrgangsstufe

Die Altersstruktur der Stichprobe stellt sich, wie in Kapitel 4.1 beschrieben, sehr heterogen dar (vgl. Abb.4). Dennoch lassen sich alle Teilnehmer entweder der elften oder zwölften Jahrgangsstufe zuordnen. Ein Vergleich dieser beiden Jahrgangsstufen wäre also möglich, bietet sich bei dieser Stichprobe allerdings nicht an, weil diejenigen Schülerinnen und Schüler, die der elften Jahrgangsstufe zugeordnet werden können identisch mit der Stichprobe sind, die einen bilingualen Zug besuchen. Die Unterschiede, die zwischen den Jahrgangsstufen ermittelt würden, wären daher nicht eindeutig auf den Jahrgangsunterschied zurückzuführen. Des Weiteren bietet sich der Vergleich benachbarter Jahrgangsstufen weniger an, weil die individuellen Entwicklungsstadien der Schülerinnen und Schüler generelle Tendenzen verschleiern könnten, die bei einem Vergleich zwischen Schülerinnen und Schülern der zehnten und zwölften Jahrgangsstufe deutlich dargestellt werden könnten.

Aus diesem Grund wurde die Entscheidung getroffen, die Stichprobe in drei Gruppen aufzuteilen, und zwar die 15-16-Jährigen, die 17-18-Jährigen und die 19-21-Jährigen. Zwar unterscheidet sich die Stichprobengröße der drei Gruppen beträchtlich, allerdings erlaubt die Unterteilung eine sinnvolle Evaluierung des Einflusses des Alters eines Schülers auf die Bewertungskompetenz. In Tabelle 8 sind die durchschnittlichen Ergebnisse der drei Altersgruppen bezüglich des Kognitionswertes und des Kompetenzwertes dargestellt.

Tabelle 8: Durchschnittliche Kognitions- und Kompetenzwerte für die drei Altersgruppen

	15-16 Jährige (n=23)	17-18 Jährige (n=136)	19-21 Jährige (n=10)
Ø Kognitionswert	7,48	7,81	7,1
Ø Kompetenzwert	18	21,18	18

Es fällt auf, dass die Gruppe der 17-18-Jährigen sowohl den höchsten Kognitionswert, als auch den höchsten Kompetenzwert aufweist. Den geringsten Kognitionswert weist die Gruppe der 19-21-Jährigen auf, während der Kompetenzwert der Gruppe 1 und 3 gleich ist. Wäre die Bewertungskompetenz von den beiden Variablen kognitive Leistungsfähigkeit und Alter abhängig, so wären bei der Gruppe der 19-21-Jährigen höhere Werte zu erwarten gewesen. Das Ausbleiben dieser Werte liefert einen ersten Hinweis, dass das Alter keinen, oder nur einen geringen Einfluss auf die Bewertungskompetenz hat. Bei dieser Aussage muss allerdings immer bedacht werden, dass die deutlichen Unterschiede in der Gruppengröße auch einen entscheidenden Einfluss auf die Durchschnittswerte haben.

In den Tabellen 9 bis 12 ist die prozentuale Verteilung der Stichproben, aufgeteilt nach den drei Altersgruppen, dargestellt. Der jeweils höchste prozentuale Anteil der Stichprobe an einem Kompetenzniveau in einer Teilaufgabe, bzw. bei einer Teilkompetenz der Bewertungskompetenz ist markiert, um eine bessere Vergleichbarkeit zu gewährleisten. Generell lässt sich eine gewisse Regelmäßigkeit dahingehend feststellen, dass bei den meisten Teilkompetenzen in allen drei Altersgruppen der Hauptteil der Stichprobe auf einem Niveau liegt. Allerdings gibt es bei einigen Teilkompetenzen große Unterschiede, die im Folgenden dargestellt werden. So unterscheiden sich die Prozentwerte bei den Maxima voneinander, was

folglich in einer anderen Verteilung der Stichprobe auf die anderen Niveaus resultiert.

Bei den Aufgaben 1A und 1B sind die Werte aller drei Altersgruppen vergleichbar (vgl. Tab. 9). Die meisten Schülerinnen und Schüler erreichen das Niveau I. Diese Ergebnisse entsprechen denen, die schon in Kapitel 5.1.3 dargestellt wurden.

Tabelle 9: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1A, 1B, 1C, 1D; höchster Wert markiert. Aufteilung nach Altersgruppen: G1=15-16 Jahre(n=23); G2=17-18 Jahre (n=136); G3=19-21 Jahre (n=10)

	Wahrnehmen und Bewusstmachen moralischer Relevanz			Beurteilen			Ethisches Basiswissen			Urteilen / Argumentieren		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Niv.	1A			1B			1C			1D		
0	0,0	0,8	10,0	8,7	9,6	0,0	30,4	16,2	10,0	0,0	2,9	10,0
I	78,3	83,8	70,0	65,2	58,1	60,0	30,4	38,9	40,0	26,1	12,5	20,0
II	21,7	11,0	20,0	26,1	32,3	40,0	26,1	34,6	30,0	39,1	36,8	40,0
III	0,0	4,4	0,0	---	---	---	13,1	10,3	20,0	34,8	47,8	30,0

Bei Aufgabe 1C fällt besonders der hohe Niveau-0-Wert von G1 ins Auge. Der Anteil der Schülerantworten mit Niveau 0 ist mit 30,4% genauso hoch wie der mit Niveau I, bei dem auch die Gruppen G2 und G3 ihr Maximum haben, allerdings weisen diese beiden Gruppen folglich einen deutlich höheren Wert bei Niveau II auf. Bei Aufgabe 1D hat die Gruppe 2 ihr Maximum mit 47,8% bei Niveau III, während bei G1 und G3 mit 39,1% bzw. 40,0% Niveau II am häufigsten erreicht wurde. Diese Teilaufgabe zeigt deutlich, dass die Werte der drei Gruppen bei Niveau II eng beieinander liegen, während der Wert von G2 auf Niveau III deutlich höher ist als der der anderen beiden Gruppen. Gleichzeitig ist der Niveau I-Wert von G2 entsprechend niedriger als bei den anderen beiden Gruppen.

Tabelle 10: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1E, 4B, 5A, 5B; höchster Wert markiert. Aufteilung nach Altersgruppen: G1=15-16 Jahre(n=23); G2=17-18 Jahre (n=136); G3=19-21 Jahre (n=10)

	Folgenreflexion			Perspektivwechsel			Ethisches Basiswissen			Beurteilen		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Niv.	1E			4B			5A			5B		
0	4,3	5,9	10,0	0,0	3,7	10,0	43,5	16,9	60,0	43,5	19,1	30,0
I	60,9	62,5	60,0	26,1	19,8	20,0	30,4	36,8	30,0	13,1	15,4	20,0
II	34,8	31,6	30,0	60,9	49,3	50,0	17,4	28,2	10,0	21,7	41,2	30,0
III	---	---	---	13,0	27,2	20,0	8,7	8,1	0,0	21,7	24,3	20,0

Bei den Aufgaben 1E und 4B sind die Verhältnisse wieder vergleichbar, auffällig ist allerdings, dass die Gruppe G2 bei der Teilaufgabe 4B einen deutlich höheren Wert bei Niveau III aufweist. Bei den Aufgaben 5A und 5B haben die Gruppen G1 und G3 besonders schlecht abgeschnitten, ihr Maximum liegt bei Aufgabe 5A mit 43,5% bzw. 60,0% bei Niveau 0, während gleichzeitig nur 16,9% von G2 Niveau 0 erreicht. Das Maximum von G2 liegt bei Niveau I (vgl. Tab. 10). Die Teilaufgabe 5B zeigt ähnliche Ergebnisse, allerdings sind die Leistungen der Gruppe G3 regelmäßig auf die vier Niveaus verteilt, was der geringen Stichprobe von n=10 geschuldet ist. Dennoch liegt ein Maximum mit 30% der Antworten bei Niveau 0, während G1 ihr Maximum mit 43,5% ebenfalls bei Niveau 0 hat. Bei dieser Teilaufgabe erreichen mit 41,2% der Gruppe G2 die Mehrheit der Schüler das Niveau II. Insgesamt fällt immer wieder auf, dass die Gruppe G2 die höchsten prozentualen Werte bei Niveau III, bzw. bei der Wertung FULL CREDIT aufweist.

Tabelle 11: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5C, 5D, 5E, 5F; höchster Wert markiert. Aufteilung nach Altersgruppen: G1=15-16 Jahre(n=23); G2=17-18 Jahre (n=136); G3=19-21 Jahre (n=10)

	Argumentieren			Wahrnehmen und Bewusstmachen moralischer Relevanz			Perspektivwechsel			Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
Niv.	5C			5D			5E			5F		
0	13,0	10,3	20,0	17,4	11,0	40,0	26,1	11,8	20,0	17,4	15,4	40,0
I	60,9	41,2	40,0	21,7	27,2	20,0	34,8	33,8	10,0	39,1	24,3	10,0
II	26,1	48,5	40,0	43,5	35,3	40,0	30,4	47,1	70,0	21,7	29,4	30,0
III	---	---	---	17,4	26,5	0,0	8,7	7,3	0,0	21,7	30,9	20,0

Bei den Aufgaben 5C und 5D treten bei der Gruppe G3 erneut zwei Maxima auf, was eine Auswertung der Ergebnisse für diese Gruppe erschwert. Bei Aufgabe 5C zeigt G1 sein Maximum bei Niveau I. Der Wert ist mit 60,9% deutlich höher als der Niveau-I-Wert der Gruppen G2 und G3, der jeweils um 40% liegt. Das Maximum der Gruppen G2 und G3 liegt folglich auf Niveau II. Das entspricht der insgesamt hohen Teilkompetenz „Argumentieren“, die bereits in Kapitel 5.1.3 festgestellt wurde. Hohe Schwankungen gibt es bei Aufgabe 5F. Während die Gruppen G1 und G3 ihr Maximum bei Niveau I (39,1%) und Niveau 0 (40,0%) haben, liegt das Maximum der Gruppe G2 mit 30,9% bei Niveau III (vgl. Tab. 11).

Tabelle 12: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungseinheiten (FC= FULL CREDIT, HC=HALF CREDIT, NC=NO CREDIT) von Bewertungskompetenz, Aufgabe 2, 3, 4A; höchster Wert markiert. Aufteilung nach Altersgruppen: G1=15-16 Jahre(n=23); G2=17-18 Jahre (n=136); G3=19-21 Jahre (n=10)

	Ethisches Basiswissen			Wahrnehmen und Bewusstmachen moralischer Relevanz			Urteilen		
	G1	G2	G3	G1	G2	G3	G1	G2	G3
	2			3			4A		
FC	4,3	29,4	0,0	0,0	3,0	0,0	69,6	74,3	80,0
HC	26,1	26,4	60,0	52,2	67,6	70,0	---	---	---
NC	69,6	44,2	40,0	47,8	29,4	30,0	30,4	25,7	20,0

Bei den Aufgaben 2, 3 und 4A liegen wieder ausgeglichene Verhältnisse vor. Die große Mehrheit aller drei Gruppen erreicht bei Aufgabe 2 die Wertung „NO CREDIT“, nur G3 weist einen Ausreißer nach oben auf, der aber wahrscheinlich durch die geringe Stichprobengröße zu begründen ist. Bei Aufgabe 3 liegen die Maxima jeweils bei der Wertung „HALF CREDIT“ und weisen Werte von 52,2% (G1), 67,6% (G2) und 70,0% (G3) auf (vgl. Tab. 12). Bei Aufgabe 4 liegen die Maxima wiederum durchgehend bei der Wertung „FULL CREDIT“

Betrachtet man die prozentualen Verteilungen der Schülerantworten auf die Niveaus von Bewertungskompetenz kann man feststellen, dass die Trends aus Tabelle 8, sowie aus Kapitel 5.1.3 bestätigt werden. Schülerinnen und Schüler der Gruppe G2 erreichen drei von 15 Teilaufgaben eine höhere Niveaustufe als die beiden anderen Gruppen. In keiner der Aufgaben erreicht eine andere Altersgruppe ein höheres Kompetenzniveau. Besonders hohe Werte werden insgesamt bei den Teilkompetenzen „Argumentieren“ und „Urteilen“ erreicht. Auch bei den Teilkompetenzen „Perspektivwechsel“ und „Beurteilen“ sind die Ergebnisse zufriedenstellend, während besonders bei der Teilkompetenz „Wahrnehmen und Bewusstmachen ethischer Relevanz“ und „Ethisches Basiswissen“ Schwächen auftreten. Bei den anderen Teilkompetenzen unterscheiden sich die Ergebnisse innerhalb der Altersgruppen. Insgesamt zeigen die prozentualen Ergebnisse der drei Gruppen, dass die Altersgruppe der 17-18-Jährigen die besten Ergebnisse erzielt.

Abschließend kann man aus dem Kognitionswert und dem Kompetenzwert ein Punktdiagramm erstellen, das die Korrelation abbildet. Die Ausgleichsgerade, die über die Punktwolke gelegt wird, verdeutlicht dabei graphisch die Korrelation und stellt eine Trendlinie dar. Zwar hat der Wert der Steigung der Gerade keinen direkten Einfluss auf den Korrelationswert r , allerdings verhalten sich beide Werte

proportional. Je steiler die Gerade, desto höher ist auch die Korrelation. Zur Darstellung der Ergebnisse der unterschiedlichen Altersgruppen werden drei Punktdiagramme erstellt. Nachfolgend werden die Steigung der Trendlinie und der Korrelationskoeffizient r angegeben und die Korrelation wird bewertet.

In Abbildung 13 sind der Kognitionswert und der Kompetenzwert der Gruppe G1, d.h. der Gruppe der 15-16-Jährigen korreliert. Die Trendlinie die Steigung y .

$$y = 0,126x + 17,05$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .028$, was einer sehr schwachen Korrelation entspricht. Der Wert ist so gering, dass eigentlich davon sprechen kann, dass keine Korrelation vorliegt.

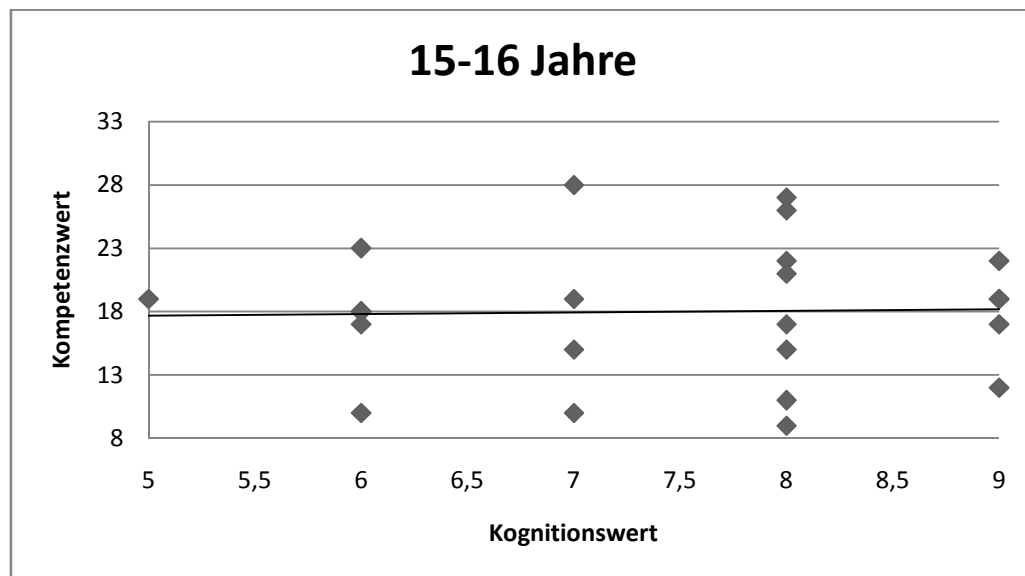


Abbildung 13: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei den 15-16-Jährigen (n=23)

In Abbildung 14 sind der Kognitionswert und der Kompetenzwert der Gruppe G2, d.h. der Gruppe der 17-18-Jährigen korreliert. Die Trendlinie hat die Steigung y .

$$y = 0,941x + 13,83$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .166$, was einer sehr schwachen Korrelation entspricht.

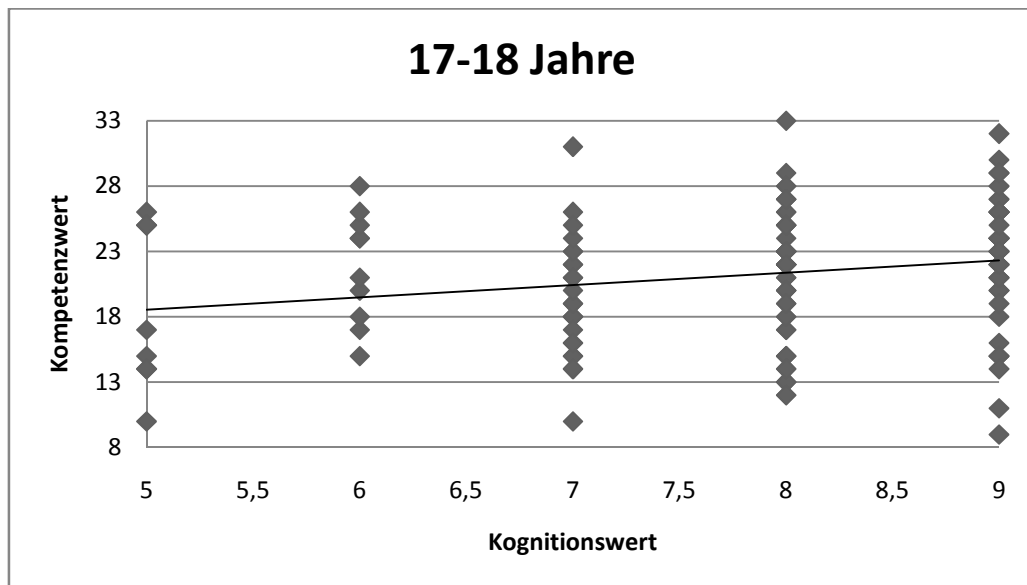


Abbildung 14: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei den 17-18-Jährigen (n=136)

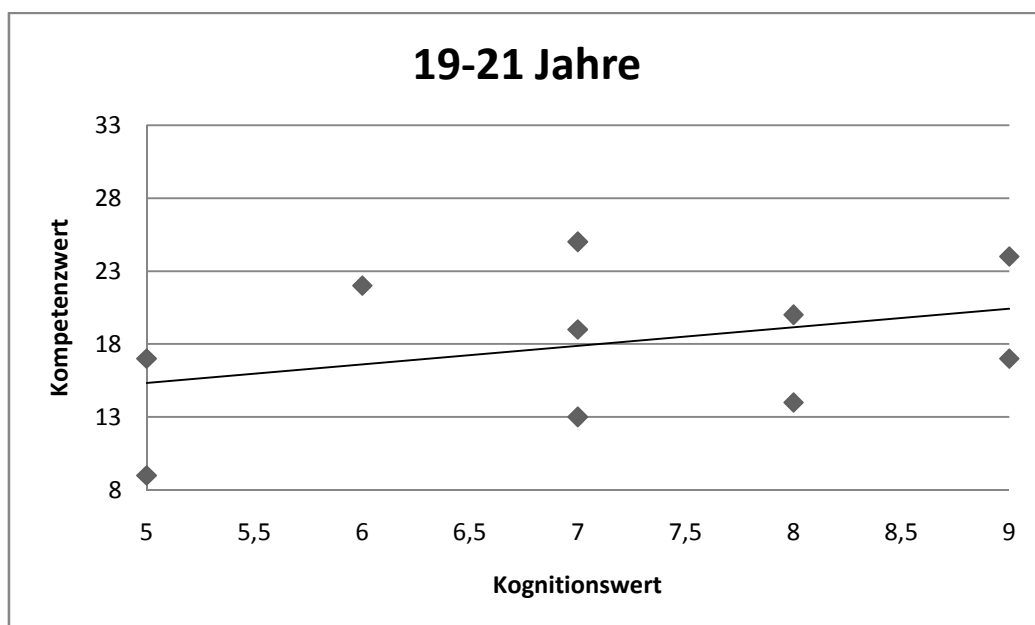


Abbildung 15: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei den 19-21-Jährigen (n=10)

In Abbildung 15 sind der Kognitionswert und der Kompetenzwert der Gruppe G3, d.h. der Gruppe der 19-21-Jährigen korreliert. Die Trendlinie hat die Steigung y .

$$y = 1,269x + 9,984$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .364$, was einer mittleren bis schwachen Korrelation entspricht. Die Korrelation der kognitiven

Leistungsfähigkeit mit der Bewertungskompetenz ist bei der Gruppe G3 ($r = .364$) am höchsten und bei G1 am niedrigsten ($r = .028$). Der Zusammenhang dieser beiden Merkmale stellt sich für G1 und G2 als schwach, für G3 als schwach bis mittelstark dar. Der Wert ist deutlich höher als bei G1 und G2, liegt dies wahrscheinlich zum Teil an der geringen Stichprobengröße.

Wie bereits bei der Beschreibung der Stichprobe angedeutet, liegt bei der untersuchten Gruppe der Sonderfall vor, dass in allen Kursen sowohl Schülerinnen und Schüler des Bildungsganges G8, als auch des älteren G9 vertreten sind (vgl. Kap. 4.1). Die Stichprobe der G9-Schüler ist mit $n=101$ größer als die G8-Gruppe ($n=60$). Acht Teilnehmer haben keine Angabe zum Bildungsgang gemacht; ihre Daten können hier nicht ausgewertet werden, weil man nicht vom Alter der Schülerinnen und Schüler automatisch auf den Bildungsgang schließen kann. Die durchschnittlichen Kognitions- und Kompetenzwerte der beiden Gruppen sind in Tabelle 13 dargestellt.

Tabelle 13: Durchschnittliche Kognitions- und Kompetenzwerte für die G8-Gruppe und die G9-Gruppe

	G8 (n=60)	G9 (n=101)
Ø Kognitionswert	7,66	7,79
Ø Kompetenzwert	20,45	21,09

Die Schülerinnen und Schüler des Bildungsganges G9 haben einen leicht höheren Kognitionswert und einen ebenfalls leicht erhöhten Kompetenzwert. Die Unterschiede sind mit 0,13 Punkten (Kognition) und 0,64 (Kompetenz) jedoch nur sehr gering. Eine deutlichen Vorteil, bzw. eine deutlich höhere Ausprägung der Bewertungskompetenz bei den G9-Schülern, durch das zusätzliche Schuljahr ist nicht zu erkennen.

Die prozentualen Ergebnisse der beiden Gruppen, die in den Tabellen 14-16 dargestellt sind, bestätigen den Trend der geringen Unterschiede. Bei vielen Teilaufgaben haben beide Gruppen ihr Maximum auf dem gleichen Niveau (Aufgaben 1A, 1B, 1C, 1E, 2, 3, 4A, 4B, 5B, 5D und 5E). Zwar unterscheiden sich die jeweiligen Prozentwerte der Gruppen dabei teilweise deutlich, allerdings ist die generelle Vergleichbarkeit gegeben.

Tabelle 14: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1 und 4B, höchster Wert markiert. Aufteilung nach Bildungsgang

	Wahrnehmen und Bewusstmachen moralischer Relevanz		Beurteilen		Ethisches Basiswissen		Urteilen / Argumentieren		Folgenreflexion		Perspektivwechsel	
	G8	G9	G8	G9	G8	G9	G8	G9	G8	G9	G8	G9
Niv.	1A		1B		1C		1D		1E		4B	
0	0,0	2,0	10,0	7,9	15,0	17,8	0,0	5,0	1,7	8,9	1,7	5,0
I	85,0	79,2	68,3	52,5	48,3	32,7	8,3	16,8	66,7	58,4	18,3	21,8
II	10,0	15,8	21,7	39,6	30,0	34,6	50,0	30,7	31,6	32,7	55,0	47,5
III	5,0	3,0	---	---	6,7	14,9	41,7	47,5	---	---	25,0	25,7

Oftmals ist der Wert beim Maximum für die G8-Gruppe höher, dafür ist der Wert des nächsthöheren Niveaus bei der G9-Gruppe höher (vgl. beispielsweise 1A und 1B; Tab. 14). Bei der Aufgabe 1D liegt das Maximum der G8-Gruppe mit 50% auf Niveau II, während das Maximum der G9-Gruppe mit 47,5% bei Niveau III liegt. Allerdings liegen mit 41,7% der G8-Schüler ähnlich viele Schüler auf Niveau III. Bei dieser Teilaufgabe stellt sich die Leistung der G9-Gruppe heterogener dar als die der G8-Gruppe.

Tabelle 15: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5, höchster Wert markiert. Aufteilung nach Bildungsgang

	Ethisches Basiswissen		Beurteilen		Argumentieren		Wahrnehmen und Bewusstmachen moralischer Relevanz		Perspektivwechsel		Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren	
	G8	G9	G8	G9	G8	G9	G8	G9	G8	G9	G8	G9
Niv.	5A		5B		5C		5D		5E		5F	
0	25,0	20,8	25,0	18,8	11,7	8,9	11,7	12,9	16,7	13,8	11,7	19,9
I	31,7	36,7	15,0	13,7	45,0	42,6	26,7	25,7	36,7	27,7	30,0	22,7
II	36,7	33,6	33,3	43,7	43,3	48,5	36,6	37,6	38,3	51,5	23,3	31,7
III	6,6	8,9	26,7	23,8	---	---	25,0	23,8	8,3	7,0	35,0	25,7

Bei den Aufgaben 5A und 5C liegen die Maxima der beiden Gruppen nicht auf dem gleichen Niveau. Bei 5A liegt das Maximum der G8-Gruppe höher, während bei 5C das Maximum der G9-Gruppe höher liegt. Dennoch sind die Werte untereinander sehr ähnlich, große Leistungsunterschiede treten hier nicht auf, die prozentualen Unterschiede haben wahrscheinlich statistische Gründe. Bei der Teilaufgabe 5F schneiden die G8-Schüler allerdings besser ab, als die G9-Schüler. 35,0% der G8-Gruppe erreichen Niveau III, was nur bei 25,7% der G9-Gruppe der Fall ist.

Tabelle 16: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungseinheiten (FC= FULL CREDIT, HC=HALF CREDIT, NC=NO CREDIT) von Bewertungskompetenz, Aufgabe 2, 3, 4A; höchster Wert markiert. Aufteilung nach Bildungsgang

	Ethisches Basiswissen		Wahrnehmen und Bewusstmachen moralischer Relevanz		Urteilen	
	G8	G9	G8	G9	G8	G9
Wertung	2		3		4A	
FC	21,7	27,7	0,0	4,0	73,3	75,2
HC	20,0	34,6	68,3	66,3	---	---
NC	58,3	37,7	31,7	29,7	26,7	24,8

Tabelle 16 zeigt, dass sich die allgemeinen Trends, nämlich große Probleme beim Lösen der Aufgaben 2 und 3 auch dann herauskristallisieren, wenn die Stichprobe anhand des Bildungsganges aufgeteilt wird. Die Werte sind für die Aufgaben 3 und 4A nahezu identisch, bei Aufgabe 2 weisen die Gruppen einen ähnliche FULL CREDIT-Wert auf, allerdings konnten ca. 15% der G9-Gruppe mehr die Wertung HALF CREDIT erreichen, was folglich in einem ebenso kleineren NO CREDIT-Wert resultiert. Insgesamt zeigen die prozentualen Ergebnisse der beiden Gruppen keine gravierenden Unterschiede was die Bewertungskompetenz angeht.

In Abbildung 16 sind der Kognitionswert und der Kompetenzwert der G8-Gruppe korreliert. Die Trendlinie hat die Steigung y.

$$y = 1,021x + 12,61$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .231$, was einer schwachen Korrelation entspricht.

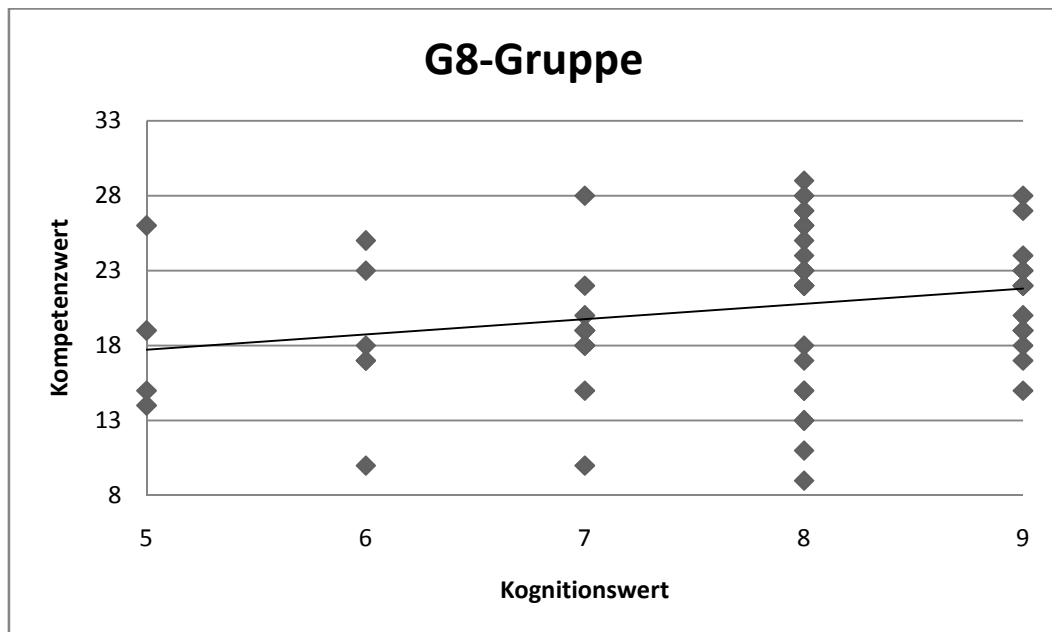


Abbildung 16: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern des gymnasialen Bildungsganges G8 (n=60)

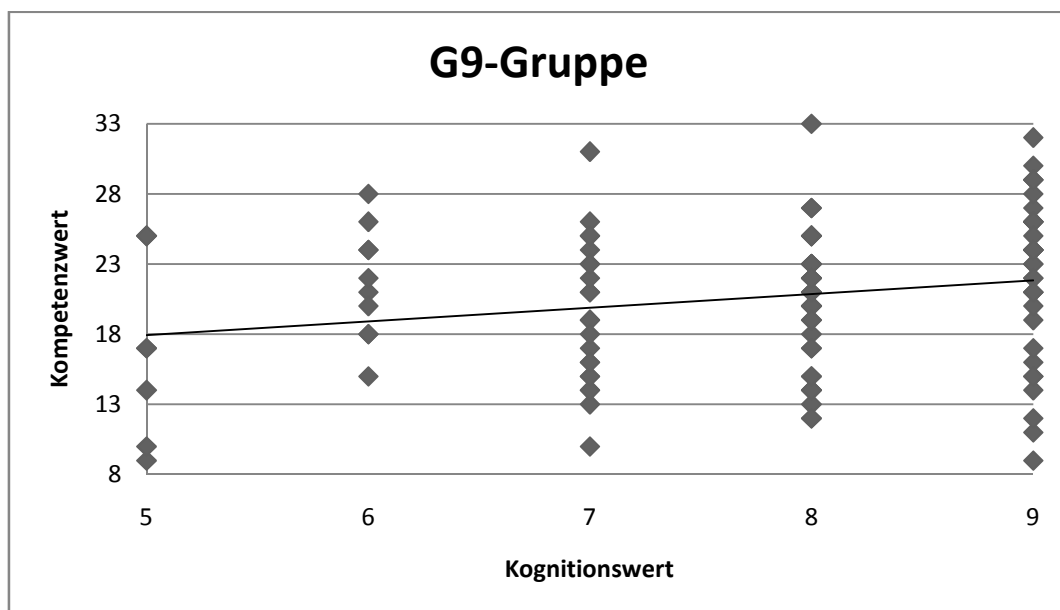


Abbildung 17: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern des gymnasialen Bildungsganges G9 (n=101)

In Abbildung 17 sind der Kognitionswert und der Kompetenzwert der G9-Gruppe korreliert. Die Trendlinie hat die Steigung y.

$$y = 0,977x + 13,04$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .218$, was einer mittleren bis schwachen Korrelation entspricht. Die Korrelation von kognitiver

Leistungsfähigkeit mit der Bewertungskompetenz ist bei der G8-Gruppe etwas höher, der Zusammenhang ist allerdings bei beiden Gruppen nur schwach.

5.2.2 Unterscheidung zwischen Biologie bilingual und Biologie auf Deutsch

Als nächstes wurde die Stichprobe nach dem besuchten Biologieunterricht aufgeteilt. Die eine Gruppe bildeten alle Schülerinnen und Schüler, die bilingualen Biologieunterricht gewählt haben. Die Größe der Stichprobe beträgt $n=105$. Diese Gruppe wird im Folgenden *bili-Gruppe* bezeichnet. Diejenigen Schülerinnen und Schüler, die Biologieunterricht auf Deutsch besuchen, bilden die zweite Gruppe ($n=64$), die folglich *mono-Gruppe* genannt wird.

Tabelle 17: Durchschnittliche Kognitions- und Kompetenzwerte für die bili-Gruppe und die mono-Gruppe

	bili-Gruppe (n=105)	mono-Gruppe (n=64)
Ø Kognitionswert	7,71	7,73
Ø Kompetenzwert	20,44	20,76

Die Durchschnittswerte liegen sehr nah beieinander. Die mono-Gruppe hat einen um 0,02 Punkte besseren Kognitionswert, während der Kompetenzwert um 0,32 Punkte besser ist als bei der bili-Gruppe. Die Durchschnittswerte suggerieren also nahezu identische kognitive Leistungsfähigkeiten der beiden Gruppen und, was für die Minimalforderung für bilingualen Unterricht besonders wichtig ist, einen nahezu identischen Kompetenzwert.

Auch bei dieser Unterteilung der Gesamtstichprobe unterscheiden sich die prozentualen Werte nicht gravierend. In Tabelle 18 liegen die Maxima außer bei Teilaufgabe 1D bei beiden Gruppen auf der gleichen Niveaustufe. Zwar unterscheiden sich die Werte der beiden Gruppen bei den Teilkompetenzen um einige Prozentpunkte, eine generelle Tendenz ist aber sichtbar. Bei der angesprochenen Teilaufgabe 1D hat die bili-Gruppe ihr Maximum bei Niveau III (47,6%), während die mono-Gruppe ihr Maximum bei Niveau II (48,4%) aufweist. Bei der Aufgabe 1 und 4B weist die bili-Gruppe durchgängig etwas höhere Niveau-III-Werte auf als die bili-Gruppe. Auch die Niveau-I-Werte der bili-Gruppe sind bei diesen Aufgaben kleiner.

Tabelle 18: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1 und 4B höchster Wert markiert. Aufteilung nach Biologieunterricht; bili=bilingualer Biologieunterricht (n=105), mono=Biologieunterricht auf Deutsch (n=64)

	Wahrnehmen und Bewusstmachen moralischer Relevanz		Beurteilen		Ethisches Basiswissen		Urteilen / Argumentieren		Folgenreflexion		Perspektivwechsel	
	bili	mono	bili	mono	bili	mono	bili	mono	bili	mono	bili	mono
Niv.	1A		1B		1C		1D		1E		4B	
0	2,0	0	6,7	12,5	16,2	20,3	2,0	4,7	4,7	7,8	1,0	7,8
I	80,0	85,9	59,0	59,4	35,2	42,2	20,0	6,3	58,2	68,8	19,0	23,4
II	14,2	10,9	34,3	28,1	33,4	32,8	30,4	48,4	37,1	23,4	53,3	46,9
III	3,8	3,2	---	---	15,2	4,7	47,6	40,6	---	---	26,7	21,9

Die Tabelle 19 zeigt ausgeglichene Maxima für die Teilaufgaben 5B, 5D und 5E, allerdings ist der Wert der mono-Gruppe jeweils deutlich höher. Das bedeutet, dass jeweils mehr Schülerinnen und Schüler der mono-Gruppe Niveau II und III erreichen, als dies bei der bili-Gruppe der Fall ist. Bei den Teilaufgaben 5C und 5F liegen zudem die Maxima der mono-Gruppe auf einem höheren Niveau als die der bili-Gruppe. Nur bei der Teilaufgabe 5A ist das Verhältnis umgekehrt. Während im ersten Teil des Fragebogens also die bili-Gruppe erfolgreicher abschneidet, erreichen bei Aufgabe 5 mehr Schülerinnen und Schüler der mono-Gruppe ein höheres Niveau. Bei Aufgabe 5 ist der Niveau-0-Wert der bili-Gruppe fast durchgängig höher als der der mono-Gruppe, während diese oft einen höheren Niveau-III-Wert hat. Die tendenziellen Ergebnisse von Aufgabe 1 und 4B sind hier also umgedreht.

Tabelle 19: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5; höchster Wert markiert. Aufteilung nach Biologieunterricht; bili=bilingualer Biologieunterricht (n=105), mono=Biologieunterricht auf Deutsch (n=64)

	Ethisches Basiswissen		Beurteilen		Argumentieren		Wahrnehmen und Bewusstmachen moralischer Relevanz		Perspektivwechsel		Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren	
	bili	mono	bili	mono	bili	mono	bili	mono	bili	mono	bili	mono
Niv.	5A		5B		5C		5D		5E		5F	
0	21,9	25,0	26,7	17,2	14,3	6,3	15,2	10,9	19,0	6,3	19,0	14,1
I	33,3	39,1	17,1	12,5	44,7	42,2	25,7	26,6	32,4	32,8	28,6	20,3
II	35,3	31,2	35,2	42,2	41,0	51,5	34,3	40,6	42,9	51,5	26,7	31,3
III	9,5	4,7	21,0	28,1	---	---	24,8	21,9	5,7	9,4	25,7	34,3

In Tabelle 20 sind die Ergebnisse für die Teilaufgaben 2, 3 und 4B zusammengefasst. Auch hier liegen wieder ausgeglichene Verhältnisse vor. Die große Mehrheit beider Gruppen erreicht bei Aufgabe 2 die Wertung „NO CREDIT“. Bei Aufgabe 3 liegen die Maxima jeweils bei der Wertung „HALF CREDIT“ und weisen Werte von 61,0% (bili-Gruppe), und 71,9% (mono-Gruppe) auf (vgl. Tab. 20). Bei Aufgabe 4 liegen die Maxima wiederum durchgehend bei der Wertung „FULL CREDIT“. Die NO CREDIT-Werte der bili-Gruppe sind etwas höher als die der mono-Gruppe.

Tabelle 20: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungen der Bewertungskompetenz, Aufgabe 2, 3 und 4A; höchster Wert markiert. Aufteilung nach Biologieunterricht; bili=bilingualer Biologieunterricht (n=105), mono=Biologieunterricht auf Deutsch (n=64). FC=FULL CREDIT, HC=HALF CREDIT, NC=NO CREDIT

	Ethisches Basiswissen		Wahrnehmen und Bewusstmachen moralischer Relevanz		Urteilen	
	bili	mono	bili	mono	Bili	mono
Wertung	2		3		4B	
FC	26,7	20,3	2,8	3,1	72,4	76,6
HC	23,8	35,9	61,0	71,9	---	---
NC	49,5	43,8	36,2	25,0	27,6	23,4

Insgesamt werden auch bei dieser Gruppeneinteilung die allgemeinen Ergebnisse bezüglich der Performanz der Schülerinnen und Schüler in Hinblick auf die einzelnen Teilkompetenzen bestätigt (vgl. Kap. 5.1.3).

In Abbildung 18 sind der Kognitionswert und der Kompetenzwert der bili-Gruppe korreliert. Die Trendlinie hat die Steigung y.

$$y = 1,557x + 8,42$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .349$, was einer mittleren bis schwachen Korrelation entspricht.

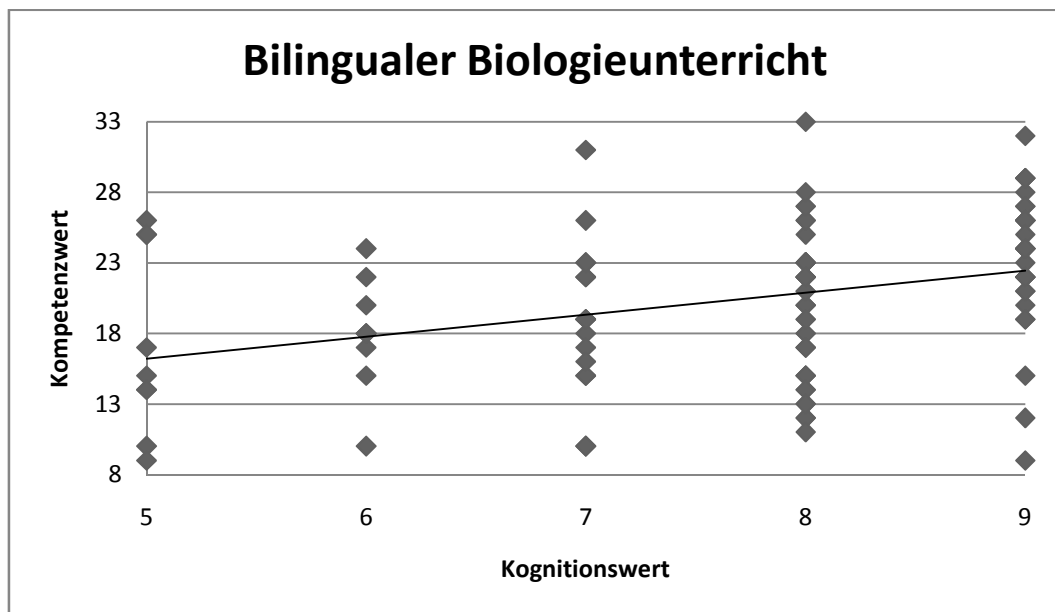


Abbildung 18: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern mit bilinguaem Biologieunterricht (n=105)

In Abbildung 19 sind der Kognitionswert und der Kompetenzwert der mono-Gruppe korreliert. Die Trendlinie hat die Steigung y.

$$y = -0,156x + 21,97$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = -.035$, was einer sehr schwachen Korrelation entspricht. Der Wert ist so gering, dass man fast von keiner Korrelation sprechen kann.

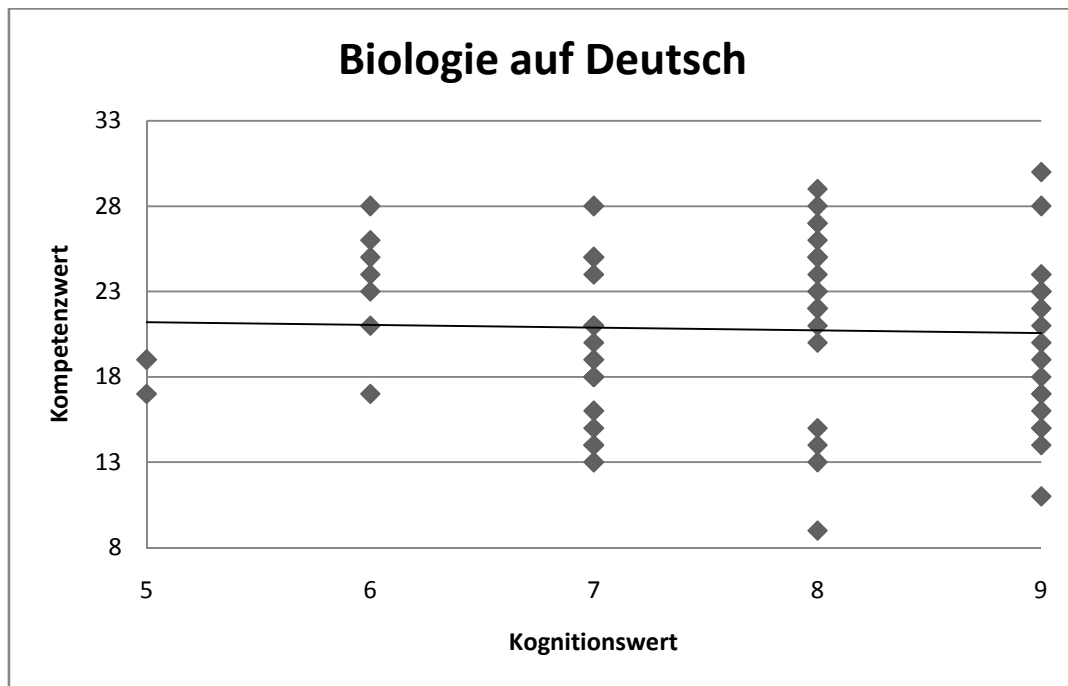


Abbildung 19: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern mit Biologie auf Deutsch (n=64)

Man könnte die beiden Gruppen erneut aufteilen und nach der jeweiligen Altersstruktur untersuchen. Dieser ausführliche Vergleich mit prozentualer Verteilung der Schülerleistungen auf die Teilaufgaben wurde allerdings nicht mit der gleichen Alterseinteilung durchgeführt wie in Kapitel 5.2.1, da bei jener Einteilung sehr kleine Stichproben entstanden wären. Die Gruppe „Biologie auf Deutsch, 19-21-Jährige“ würde nur vier Werte umfassen, die Stichprobe „Biologie bilingual, 19-21-Jährige“ wäre mit n=6 nur unwesentlich größer. Daher wurde die Stichprobe nur in zwei Untergruppen aufgeteilt, nämlich die 15-17-Jährigen und die 18-21-Jährigen. Aus dieser Einteilung resultieren vier Untergruppen: bilingual unterrichtete 15-17-Jährige (bG1), bilingual unterrichtete 18-21-Jährige (bG2), monolingual unterrichtete 15-17-Jährige (mG1) und monolingual unterrichtete 18-21-Jährige (mG2).

Zuerst werden erneut die Durchschnittswerte der vier Gruppen verglichen, um erste Erkenntnisse über Unterschiede in kognitiver Leistungsfähigkeit und Bewertungskompetenz zu erkennen (s. Tab. 21).

Tabelle 21: Durchschnittliche Kognitions- und Kompetenzwerte für die Gruppen bG1, bG2, mG1 und mG2

	bG1(n=66)	bG2 (n=39)	mG1 (n=36)	mG2 (n=28)
Ø Kognitionswert	7,62	7,87	7,72	7,75
Ø Kompetenzwert	19,89	21,36	21,08	20,35

Ein positiver Nebeneffekt dieser Einteilung sind Stichprobengrößen, die sich immer noch unterscheiden, deren Unterscheidung jedoch nicht so stark ausgeprägt ist, dass aussagekräftige Daten ermittelt werden können.

Beim Vergleich der Werte fällt auf, dass sich der durchschnittliche Kognitionswert der vier Gruppen nur unwesentlich voneinander unterscheidet. Der Wert der jeweils älteren Stichprobe ist leicht erhöht (0,25 Punkte zwischen bG1 und bG2; 0,02 Punkte zwischen mG1 und mG2). Ein größerer Unterschied findet sich beim Kompetenzwert zwischen den Gruppen bG1 und bG2. Die ältere Gruppe hat einen um 1,47 Punkte höheren Wert. Die Werte von mG1 und mG2 unterscheiden sich um 0,73 Punkte, allerdings weist hier die jüngere Gruppe den höheren Wert auf. Den höchsten Kognitionswert hat die Gruppe bG2 mit 21,36, den niedrigsten hat die Gruppe bG1 mit 19,89. Auch bei den durchschnittlichen Kognitionswerten haben diese beiden Gruppen den höchsten und den niedrigsten Wert.

Aus Gründen der Übersichtlichkeit werden im Folgenden bei den Prozenttabellen nur diejenigen Teilaufgaben abgebildet, in denen es Maxima-Unterschiede innerhalb der Gruppen gibt. Nicht explizit dargestellt sind die Aufgaben 1A, 1E, 3, 4A, 4B, 5B, 5D und 5E. Bei diesen Teilaufgaben haben alle vier Gruppen ihr Maximum auf dem gleichen Niveau. Zwar gibt es bei den Werten selbst einige kleinere Unterschiede, auch im Hinblick auf die restliche prozentuale Verteilung auf die anderen Niveaus von Bewertungskompetenz. Diese Unterschiede sind allerdings nicht allzu signifikant.

Tabelle 22: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 1B-1D; höchster Wert markiert. Aufteilung nach Biologieunterricht und Alter

	Beurteilen				Ethisches Basiswissen				Urteilen / Argumentieren			
	bG1	bG2	mG1	mG2	bG1	bG2	mG1	mG2	bG1	bG2	mG1	mG2
Niv.	1B				1C				1D			
0	7,6	5,1	16,7	7,1	19,7	10,2	11,1	32,1	0,0	5,1	2,7	7,1
I	66,7	46,1	69,4	46,4	37,9	30,8	44,4	39,3	15,1	28,2	8,3	3,6
II	25,7	48,8	13,9	46,5	30,3	38,5	36,2	28,6	36,4	20,5	52,8	42,9
III	---	---	---	---	12,1	20,5	8,3	0,0	48,5	46,2	36,1	46,4

Bei der Aufgabe 1B haben die beiden älteren Gruppen (bG2 und mG2) ihr Maximum auf Niveau II, während die jüngeren Gruppen ihr Maximum auf Niveau I haben. Die Prozentwerte der Maxima von bG2 und mG2 sind aber nur etwas höher als die Werte bei Niveau I (vgl. Tab. 22). Diese beiden Stichprobengruppen sind regelmäßig auf

Niveau I und II verteilt, während die Maxima der beiden anderen Gruppen deutlich höhere Prozentwerte aufweisen als bei den anderen Niveaus.

Bei Aufgabe 1C liegen die Maxima der Gruppen bG1, mG1 und mG2 bei Niveau I, nur bG2 hat ihr Maximum bei Niveau II. Auffällig ist außerdem ein sehr hoher Wert der Gruppe mG2 bei Niveau 0 und daraus resultierend der Nullwert auf Niveau III. Insgesamt liegen deutlich weniger Schülerantworten von mG1 und mG2 bei Niveau III.

Bei Aufgabe 1D haben die beiden bilingualen Gruppen deutlich höhere Werte bei Niveau I, als die monolingualen Gruppen. Die Maxima der Gruppe bG1, bG2 und mG2 liegt auf Niveau III, das von mG1 liegt auf Niveau II. Der mG1-Wert für Niveau II ist außerdem deutlich höher als der für Niveau III.

Tabelle 23: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5A, 5C und 5F; höchster Wert markiert. Aufteilung nach Biologieunterricht und Alter

	Ethisches Basiswissen				Argumentieren				Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren			
	bG1	bG2	mG1	mG2	bG1	bG2	mG1	mG2	bG1	bG2	mG1	mG2
Niv.	5A				5C				5F			
0	24,2	17,9	19,5	32,1	18,2	7,7	2,8	10,8	16,7	23,1	13,9	14,3
I	36,4	28,2	33,3	46,5	50,0	35,9	41,7	42,8	27,2	30,8	22,2	17,9
II	30,3	43,6	38,9	21,4	31,8	54,4	55,5	46,4	28,9	23,1	27,8	35,7
III	9,1	10,3	8,3	0,0	---	---	---	---	27,2	23,0	26,1	32,1

Bei Aufgabe 5A werden die Ergebnisse der Durchschnittswerte reproduziert. Die Maxima der Gruppen bG2 und mG1 liegen auf einem höheren Niveau (II) als die der Gruppen bG1 und mG2 (I). Diese beiden Gruppen haben außerdem einen höheren Niveau-0-Wert.

Bei 5C sieht das Ergebnis ganz ähnlich aus, allerdings hat hier auch mG2 sein Maximum bei Niveau II. Die Gruppe bG1 hat einen hohen Niveau-0-Wert und ihr Maximum mit 50,0% bei Niveau I.

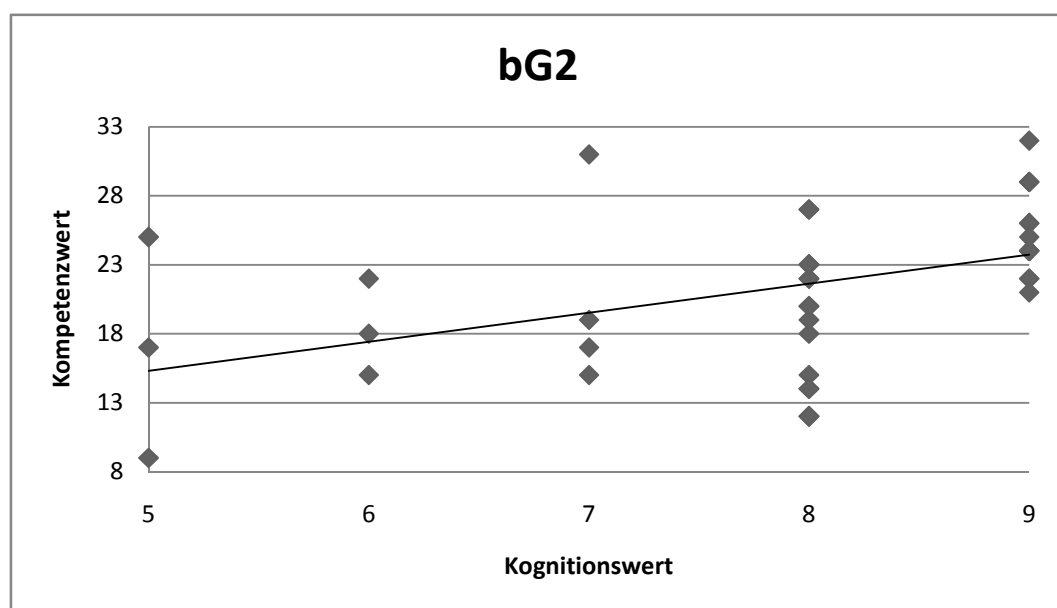
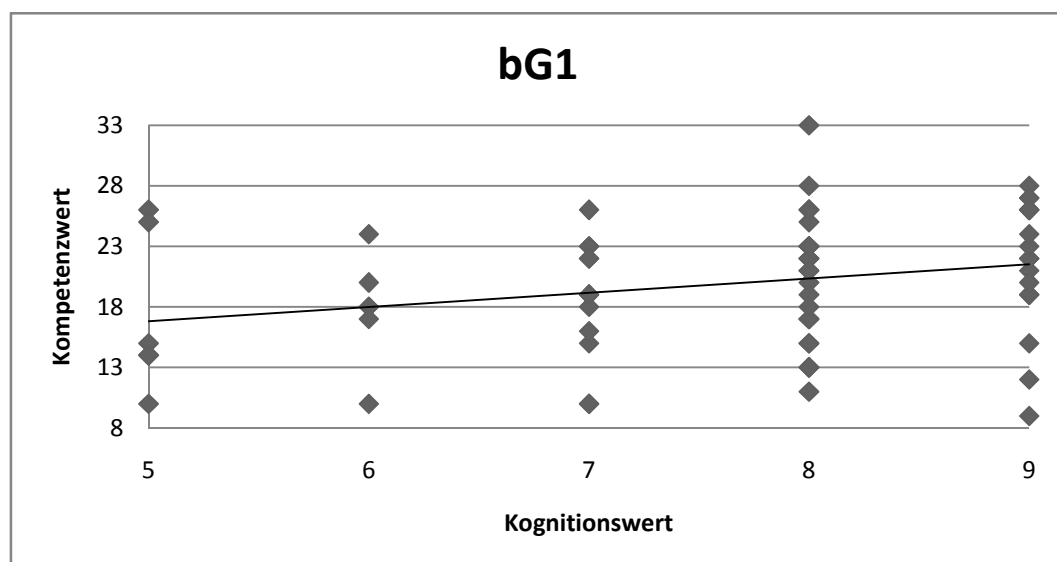
Bei Aufgabe 5F sind die Niveau-0-Werte der Gruppen vergleichbar, nur Gruppe bG2 hat einen höheren Wert. Folglich liegt auch das Maximum dieser Gruppe auf Niveau I und damit niedriger als das der anderen Gruppen. Die Niveau-III-Werte sind dann allerdings wieder ähnlich hoch. Nur Gruppe mG2 hat mit 32,1% auf Niveau III einen höheren Wert als die anderen Gruppen.

Bei Aufgabe 2 hat wiederum Gruppe mG2 den besten Wert. Während die anderen drei Gruppen ihr Maximum bei NO CREDIT haben, haben 57,2% von mG2 die Wertung HALF CREDIT erreicht (vgl. Tab. 24).

Tabelle 24: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungen Bewertungskompetenz, Aufgabe 2; höchster Wert markiert. Aufteilung nach Biologieunterricht und Alter

	Ethisches Basiswissen			
	bG1	bG2	mG1	mG2
Niveau	2			
FC	24,2	30,8	19,4	21,4
HC	19,7	30,8	19,4	57,2
NC	56,1	38,4	61,2	21,4

Auch für diese Gruppeneinteilung lassen sich Punktdiagramme darstellen.



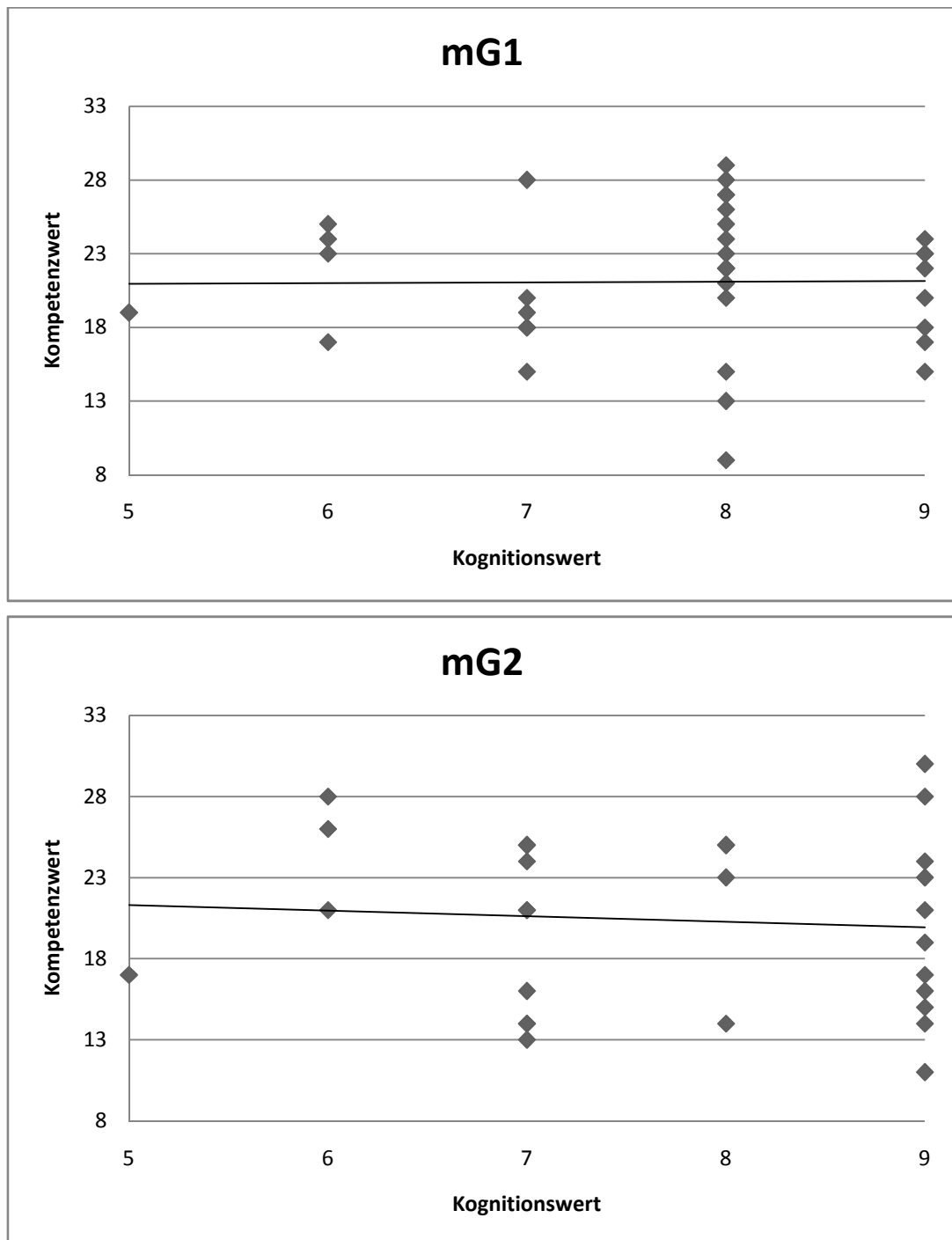


Abbildung 20: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern nach Unterrichtssprache und Alter.

Wie die Trendlinien verdeutlichen, besteht die höchste Korrelation zwischen Kognitionswert und Kompetenzwert bei der Gruppe bG2. Die Korrelationswerte sind in Tabelle 25 dargestellt. Nur für bG2 liegt eine mittlere Korrelation vor, bei den anderen Gruppen liegen schwache (bG1) und sehr schwache (mG1 und mG2) vor.

Tabelle 25: Korrelationswerte der Gruppen bG1, bG2, mG1 und mG2

	bG1	bG2	mG1	mG2
Korrelation	r = .268	r = .460	r = .011	r = -.078

5.2.3 Unterscheidung zwischen bilingualem Zug und bilingualem Modul

Für die dritte Unterfragestellung werden die Leistungen bezüglich der Bewertungskompetenz von Schülerinnen und Schülern verglichen, die bilingualen Biologieunterricht im Zuge eines bilingualen Zugs belegen und solchen, die bilinguale Module im Fach Biologie besuchen. Diese beiden Gruppen werden im Folgenden mit den Kürzeln Zug-Gruppe und Modul-Gruppe verglichen. Wie bereits in Kapitel 4.1 beschrieben, werden zu der Zug-Gruppe auch diejenigen Schülerinnen und Schüler gezählt, die neben Biologie noch in anderen Fächern bilingualen Unterricht besuchen. Diese Anpassung wird dadurch notwendig, dass nur eine der vier Schulen bilinguale Züge anbietet. Einige der entscheidenden Vorteile von bilingualen Zügen, nämlich die vermehrte Auseinandersetzung mit der englischen Fachsprache in nahezu immersionsähnlichen Situationen, ist auch dann gegeben, wenn ein Schüler mehrere bilinguale Fächer belegt, selbst dann, wenn diese Belegung nicht in Form eines Zuges organisiert ist. Mit dieser Anpassung der Stichprobe, ergeben sich zwei nahezu gleich große Unterstichproben. Die Ergebnisse der Zug-Gruppe (n=51) werden im Folgenden denen der Modul-Gruppe (n=53) gegenübergestellt.

Tabelle 26: Durchschnittliche Kognitions- und Kompetenzwerte für die Zug-Gruppe und die Modul-Gruppe

	Zug-Gruppe (n=51)	Modul-Gruppe (n=53)
Ø Kognitionswert	7,51	7,88
Ø Kompetenzwert	18	22,68

Ein Vergleich der Durchschnittswerte (s. Tab. 26) verdeutlicht, dass zwar beide Gruppen ähnlich hohe durchschnittliche Kognitionswerte aufweisen, während der Kompetenzwert der Modulgruppe mehr als 4,5 Punkte höher ist als der Wert der Zug-Gruppe. Durchschnittlich können die Modul-Schüler also von einem ähnlichen kognitiven Ausgangspunkt eine deutlich höhere Bewertungskompetenz aufweisen.

Tabelle 27: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5; höchster Wert markiert. Aufteilung nach bilinguaalem Zug (n=51) und bilinguaalem Modul (n=53)

	Wahrnehmen und Bewusstmachen moralischer Relevanz		Beurteilen		Ethisches Basiswissen		Urteilen / Argumentieren		Folgenreflexion		Perspektivwechsel	
	Zug	Modul	Zug	Modul	Zug	Modul	Zug	Modul	Zug	Modul	Zug	Modul
Niv.	1A		1B		1C		1D		1E		4B	
0	2,0	1,9	3,9	9,4	25,5	7,5	0,0	3,8	2,0	7,5	0,0	1,9
I	80,4	79,2	64,7	54,7	31,4	39,6	25,5	13,2	70,6	45,3	23,5	15,0
II	15,6	13,2	31,4	35,9	33,3	34,0	31,4	30,2	27,4	47,2	58,9	49,1
III	2,0	5,7	---	---	9,8	18,9	43,1	52,8	---	---	17,6	34,0

Wie schon bei den ersten beiden vergleichenden Ergebnisdarstellungen treten im ersten Teil des Testheftes nur geringe Unterschiede auf (vgl. Tab. 27). Bei den Teilaufgaben 1A, 1B, 1D und 4B unterscheiden sich nur die prozentualen Ergebnisse, während die Maxima bei den jeweils gleichen Niveaus liegen. Beim ethischen Basiswissen scheint die Zug-Gruppe besser abzuschneiden, immerhin liegt nur das Maximum der Zug-Gruppe auf Niveau II (33,3%), während bei der Modulgruppe mit 39,6% die meisten Teilnehmer auf das Niveau I entfallen. Allerdings ist der Wert der Modulgruppe für Niveau II mit 34,0% etwas höher als der der Zug-Gruppe, die wiederum einen hohen Niveau-0-Wert aufweist.

Bei Aufgabe 1E treten sehr hohe Unterschiede zwischen den beiden Gruppen auf. Die meisten Schülerinnen und Schüler der Zug Gruppe (70,6%) nennen hauptsächlich Folgen für die unmittelbar betroffenen Personen des Dilemmas (Niveau I), während in der Modul-Gruppe das Maximum mit 47,2% bei Niveau II liegt und damit fast doppelt so groß ist.

Bei Aufgabe 5 setzt sich dieser Trend fort. In jeder Teilaufgabe hat die Modul-Gruppe ihr Maximum auf einem höheren Niveau als die Zug-Gruppe, allerdings beträgt der Unterschied mit Ausnahme der Aufgabe 5B immer nur eine Niveaustufe. Nichtsdestotrotz ist dieser Unterschied aussagekräftig. In allen Teilaufgaben ist der Wert der Modul-Gruppe, die das Maximum bildet, deutlich höher als der Wert der Zug-Gruppe (vgl. Tab. 28). Ein Beispiel ist die Teilaufgabe 5A: 52,8% der

Schülerantworten liegt auf Niveau II, während gleichzeitig nur 15,7% der Zug-Gruppe auf dieses Niveau kommt. Ähnlich deutliche Verhältnisse kommen bei den Teilaufgaben 5C (58,8% zu 21,6%), 5D (34,0% zu 13,7%) und 5F (30,2% zu 21,5%) vor.

Tabelle 28: Prozentualer Anteil der Stichprobe an den Niveaus von Bewertungskompetenz, Aufgabe 5; höchster Wert markiert. Aufteilung nach bilingualem Zug (n=51) und bilingualem Modul (n=53)

	Ethisches Basiswissen		Beurteilen		Argumentieren		Wahrnehmen und Bewusstmachen moralischer Relevanz		Perspektivwechsel		Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren	
	Zug	Modul	Zug	Modul	Zug	Modul	Zug	Modul	Zug	Modul	Zug	Modul
Niv.	5A		5B		5C		5D		5E		5F	
0	33,3	11,4	45,1	9,4	25,5	3,7	21,6	9,4	23,5	15,1	21,5	17,0
I	41,2	26,4	19,6	15,0	52,9	37,8	25,5	26,4	37,3	26,4	31,5	26,4
II	15,7	52,8	21,6	49,1	21,6	58,8	39,2	30,2	37,3	49,1	21,5	30,2
III	9,8	9,4	13,7	26,4	---	---	13,7	34,0	1,9	9,4	25,5	26,4

Bei der bereits angesprochenen Teilaufgabe 5B fällt sofort der hohe Niveau-0-Wert der Zug-Gruppe auf, der gegenüber dem Wert der Modul-Gruppe nahezu fünffach erhöht ist. Dies ist deswegen beachtlich, da die Teilkompetenz „Beurteilen“ allgemein eine der Teilkompetenzen von Bewertungskompetenz ist, die bei den Schülerinnen und Schülern stark ausgeprägt ist. Während dieser Wert das Maximum der Zug-Gruppe darstellt, hat die Modul-Gruppe mit 49,1% ihr Maximum bei Niveau II.

Tabelle 29: Prozentualer Anteil der Stichprobe an den CREDIT-Wertungen der Bewertungskompetenz, Aufgabe 2, 3 und 4A; höchster Wert markiert. Aufteilung nach bilingualem Zug (n=51) und bilingualem Modul (n=53)

	Ethisches Basiswissen		Wahrnehmen und Bewusstmachen moralischer Relevanz		Urteilen	
	Zug	Modul	Zug	Modul	Zug	Modul
Wertung	2		3		4	
FC	9,8	43,4	3,9	0,0	70,6	73,6
HC	25,5	22,6	51,0	71,7	---	---
NC	64,7	34,0	45,1	28,3	29,4	26,4

Bei den Aufgaben 3 und 4A sind die Verhältnisse der beiden Gruppen allerdings wieder ähnlich; zumindest die Maxima liegen bei der gleichen Bewertungseinheit. Allerdings unterscheidet sich dieser Wert bei Aufgabe 3 doch deutlich. Während 71,7% der Modulgruppe die Wertung HALF CREDIT erreichen, trifft das nur auf 51,0% der Zug-Gruppe zu. Diese hat wiederum bei der Wertung NO CREDIT einen ca. 17% höheren Wert (vgl. Tab. 29). Besonders auffällig ist der hohe Wert der Modul-Gruppe bei der Wertung FULL CREDIT zur Aufgabe 2. Mit 43,4% liegt damit das Maximum dieser Gruppe auf dieser Wertungsstufe. Damit ist die Modul-Gruppe auch die einzige der unterschiedenen Teilstichproben, die einen solch hohen Wert bei dieser Aufgabe aufweist. Die Zug-Gruppe hat ihr Maximum bei dieser Aufgabe mit 64,7% bei der Wertung NO CREDIT.

In Abbildung 18 sind der Kognitionswert und der Kompetenzwert der Zug-Gruppe korreliert. Die Trendlinie hat die Steigung y .

$$y = 1,079x + 9,89$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .271$, was einer schwachen Korrelation entspricht.

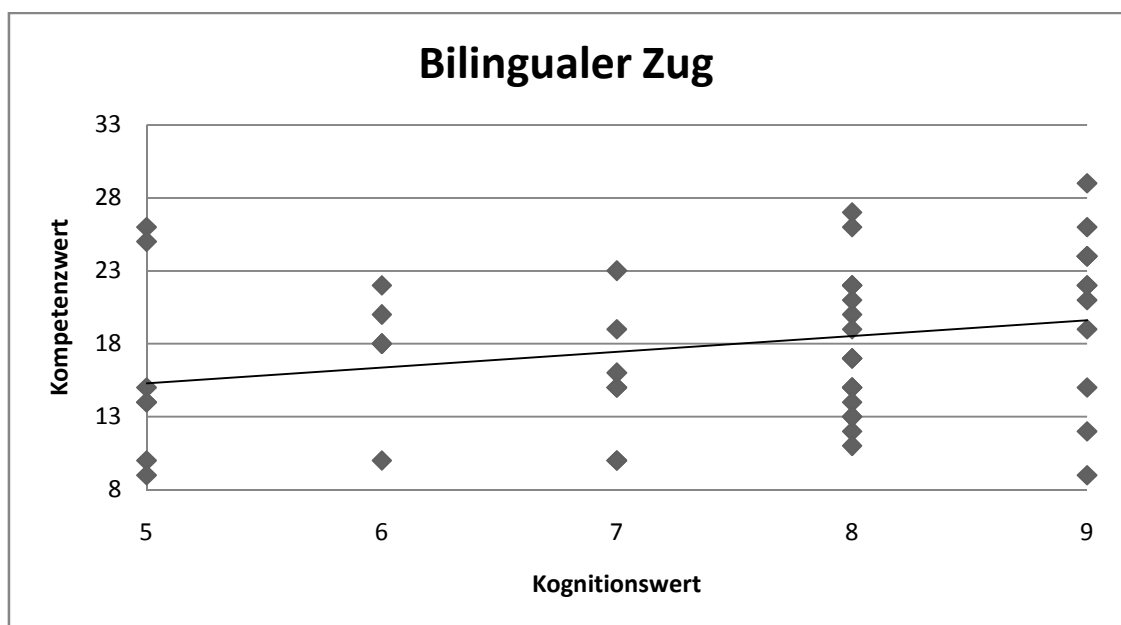


Abbildung 21: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern in einem bilingualen Zug (n=51)

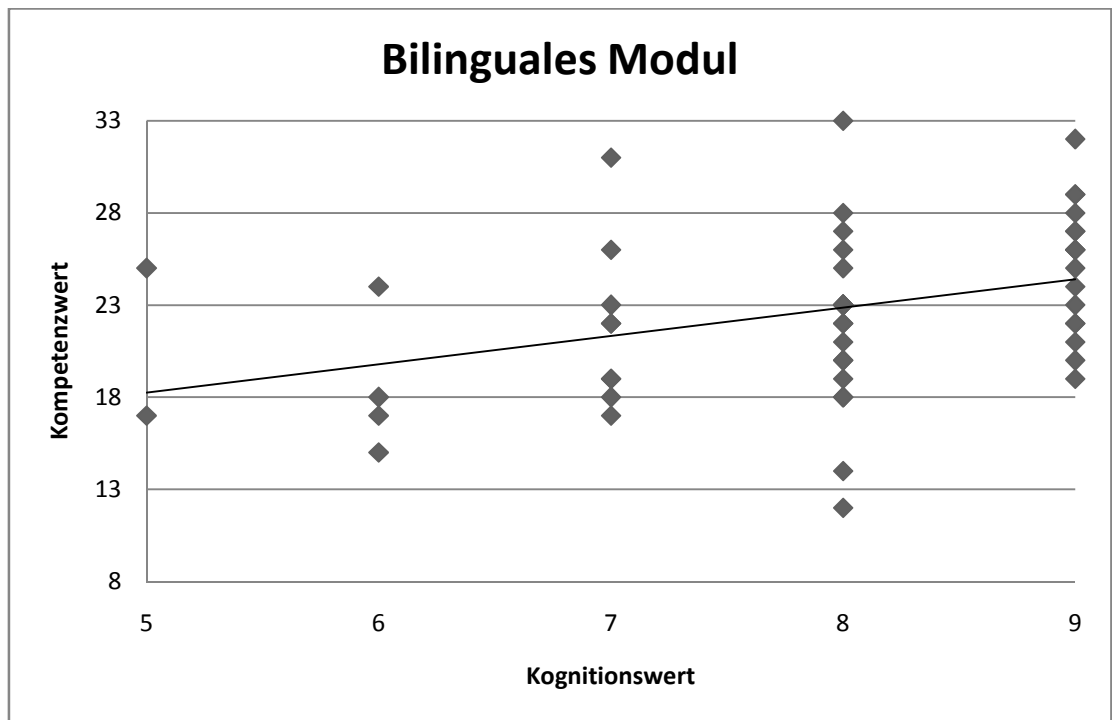


Abbildung 22: Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz bei Schülerinnen und Schülern in einem bilingualen Modul (n=53)

In Abbildung 22 sind der Kognitionswert und der Kompetenzwert der Modul-Gruppe korreliert. Die Trendlinie hat die Steigung y .

$$y = 1,535x + 10,579$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .365$, was einer mittleren bis schwachen Korrelation entspricht.

5.2.4 Zusammenfassende Ergebnisse bezüglich der Korrelation von kognitiver Leistungsfähigkeit und Ausprägung der Bewertungskompetenz.

Anschließend wurden der Kognitionswert und der Kompetenzwert Bewertungskompetenz jedes Schülers miteinander korreliert und in einem Punktdiagramm dargestellt (vgl. Abb. 23). Das Ziel dieser Datenauswertung liegt in der Untersuchung, ob diese Werte korrelieren, was in diesem Fall bedeuten würde, dass ein hoher Kognitionswert einen hohen Kompetenzwert für die Bewertungskompetenz nach sich zieht. Anders formuliert bedeutet ein hoher Korrelationswert dieser beiden Zahlen, dass eine hohe Bewertungskompetenz nur bei einer hohen kognitiven Leistungsfähigkeit erreicht werden kann.

Das Diagramm ist von mehrfachen Kombinationen der gleichen Kognitions- und Kompetenzwerte bereinigt, wodurch im Diagramm nicht alle 169 Datenpunkte

auftauchen. Die Ausgleichsgerade verdeutlicht auf graphische Weise die Korrelation. Wie schon bei den Punktdiagrammen zuvor würde eine hohe Korrelation in einer Geraden mit starker Steigung resultieren. Die deutlich flachere Gerade in Abbildung 23 verdeutlicht, dass keine starke Korrelation zwischen den beiden Werten vorliegt.

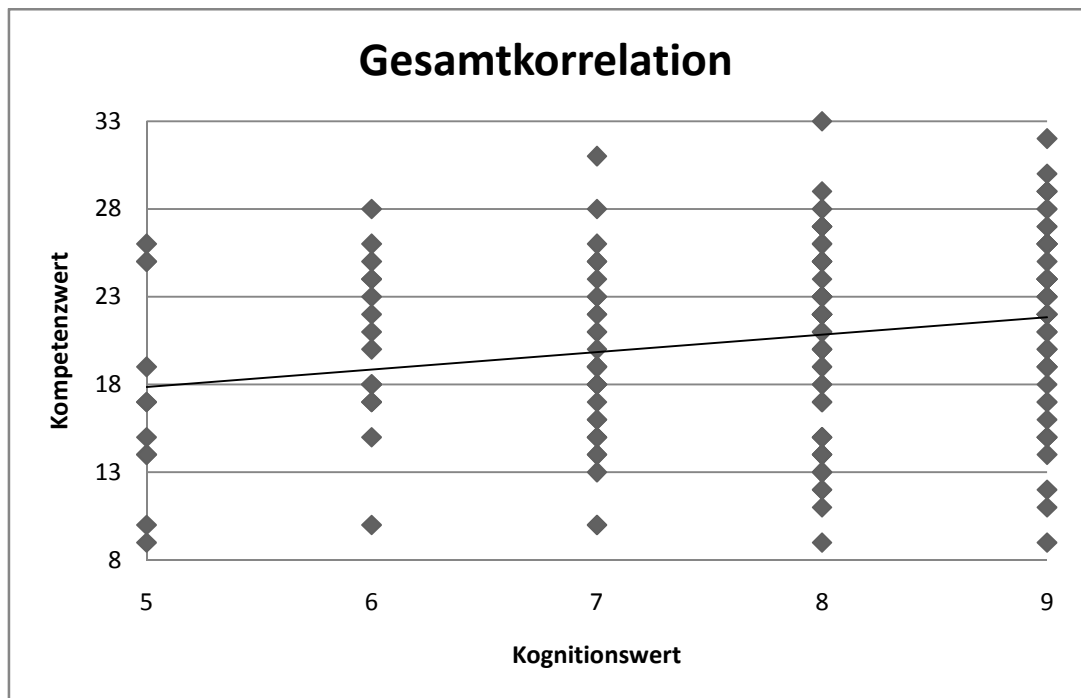


Abbildung 23: Gesamtkorrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz (n=169)

In Abbildung 23 sind der Kognitionswert und der Kompetenzwert der gesamten Stichprobe korreliert. Die Trendlinie hat die Steigung y .

$$y = 0,993x + 12,89$$

Der Korrelationskoeffizient für die Merkmalspaare der Gruppe ist $r = .223$, was einer schwachen Korrelation entspricht.

Zur besseren Übersicht werden die errechneten Korrelationswerte nochmals tabellarisch festgehalten (s. Tab. 30).

Tabelle 30: Korrelationswerte von kognitiver Leistungsfähigkeit und Bewertungskompetenz für die verschiedenen Stichproben

Stichprobe	Korrelationswert
Gesamtstichprobe (n=169)	$r = .223$
15-16-Jährige	$r = .028$
17-18-Jährige	$r = .166$
19-21-Jährige	$r = .364$
Bili-Gruppe	$r = .349$

Mono-Gruppe	$r = -.035$
bG1-Gruppe	$r = .268$
bG2-Gruppe	$r = .460$
mG1-Gruppe	$r = .011$
mG2-Gruppe	$r = -.078$
Zug-Gruppe (n=51)	$r = .271$
Modul-Gruppe (n=53)	$r = .365$

Abschließend lässt sich festhalten, dass zwischen der Bewertungskompetenz und der kognitiven Leistungsfähigkeit in den meisten Fällen nur eine schwache Korrelation besteht. Bei der bG2-Gruppe besteht mit $r = .460$ die höchste Korrelation der untersuchten Verhältnisse, allerdings liegt auch dieser Wert im unteren Bereich der mittleren Korrelation. Auch die Modul-Gruppe liegt mit $r = .365$ noch im unteren mittleren Bereich. In einigen Lehrwerken der Statistik wird bei solch geringen Werten bereits von einer geringen bzw. keiner Korrelation gesprochen. Auch der Wert der Gesamtkorrelation ist mit $r = .223$ ausgesprochen gering. Zwischen der kognitiven Leistungsfähigkeit der Schülerinnen und Schüler und ihrer Bewertungskompetenz besteht laut dieser Studie kein signifikanter Zusammenhang. Interessant ist allerdings, dass die Korrelationswerte der bili-Gruppe ($r = .349$) und der mono-Gruppe ($r = -.035$) sich am stärksten unterscheiden. Bei der bili-Gruppe hängt die kognitive Leistungsfähigkeit und die Bewertungskompetenz zehnmal stärker zusammen als bei der mono-Gruppe, allerdings liegt auch der Wert $r = .349$ noch an der Grenze zwischen schwacher und mittlerer Korrelation. Dieser Trend wird bei der Aufteilung nach Unterrichtssprache und Alter noch deutlicher. Zwischen der Gruppe mG2 ($r = -.078$) und der Gruppe bG2 ($r = .460$) besteht der größte Korrelationsunterschied. Da aber keiner der Werte eine deutliche Korrelation ausdrückt bleiben die Ergebnisse weniger aussagekräftig.

5.3. Ergebnisse Fragebogen-Items der bili-Teilnehmer

In einem dritten Teil, der nur von Teilnehmern ausgefüllt wurde, die bilingualen Biologieunterricht belegen, soll eine Selbstreflexion über ein Überforderungsempfinden im Biologieunterricht initiieren. Dazu wurden acht Statements gegeben, zu denen die Teilnehmer ihre Zustimmung auf einer Likert-Skala mit den Extrema 1 = Stimmt genau und 5 = Stimmt gar nicht angeben sollten. Abgefragt wurden die Motivation sowie eine Selbsteinschätzung der eigenen

Leistung im Vergleich zu anderen Fächern. Außerdem wurde das Überforderungsempfinden in Bezug auf Inhalt und Sprache abgefragt sowie eine

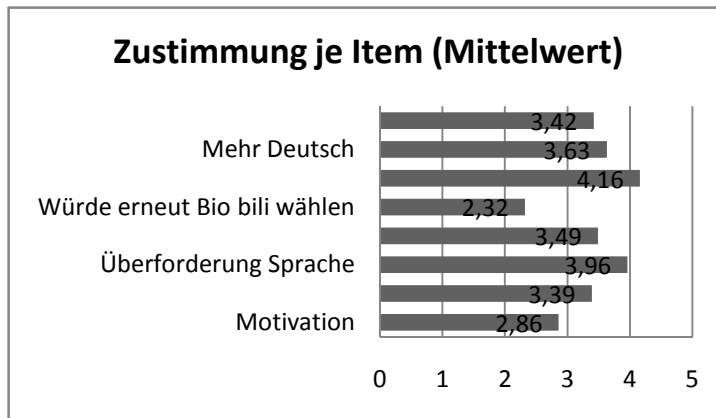


Abbildung 24: Durchschnittliche Zustimmung je Item (1= Stimmt genau; 5= Stimmt gar nicht; n=104)

Bewertung der Kurswahl und die Verwendung der deutschen Sprache. Die genauen Statements können der Testheftversion im Anhang entnommen

werden (Anhang 1). Die durchschnittlichen

Zustimmungen sind in Abbildung 24 dargestellt. Die höchste Zustimmung (2,32) erhielt die Aussage „Ich würde Biologie wieder als bilinguales Fach wählen“, während analog dazu die Aussage „Ich bereue meine Entscheidung, Biologie bilingual gewählt zu haben“ mit einem Wert von 4,16 die geringste Zustimmung erfährt. Was das Empfinden einer Überforderung angeht, identifizieren die Teilnehmer eher den Inhalt als die englische Sprache als Grund für diese. Dem Statement, dass die Überforderung sprachliche Gründe hat mit einem Wert von 3,96 die zweitgeringste Zustimmung. Teilt man die Stichprobe gemäß der ermittelten Überforderungsschwelle auf ergibt sich folgendes Bild (Abb. 25). Es fällt auf, dass in allen Schwellengruppen eher einer inhaltlichen Überforderung zugestimmt wird, wobei auch diese Werte mit 3,33 bis 3,77 weit über dem Durchschnitt der Likert-Skala liegen und damit nicht von einer Zustimmung, also einem Überforderungsempfinden ausgegangen werden kann. Bei der sprachlichen Überforderung sind die Werte sogar mit 3,55-4,13 noch näher am entgegengesetzten Pol der Zustimmung. Von Interesse ist außerdem, dass diejenigen Schüler, die die geringste Schwelle der kognitiven Überforderung haben, eine Überforderung am wenigsten wahrnehmen.

Des Weiteren wird die Selbsteinschätzung der Überforderung für verschiedene Leistungsgruppen betrachtet. In der bili-Gruppe wurden Kompetenzwerte von 9 bis 33 erreicht, was eine Einteilung der Stichprobe in 5 Gruppen erlaubt (vgl. Tab. 31). Zusammenfassend lässt sich aussagen, dass die Motivation generell mit höherem Kompetenzwert steigt, während einer sprachlichen Überforderung mit steigendem

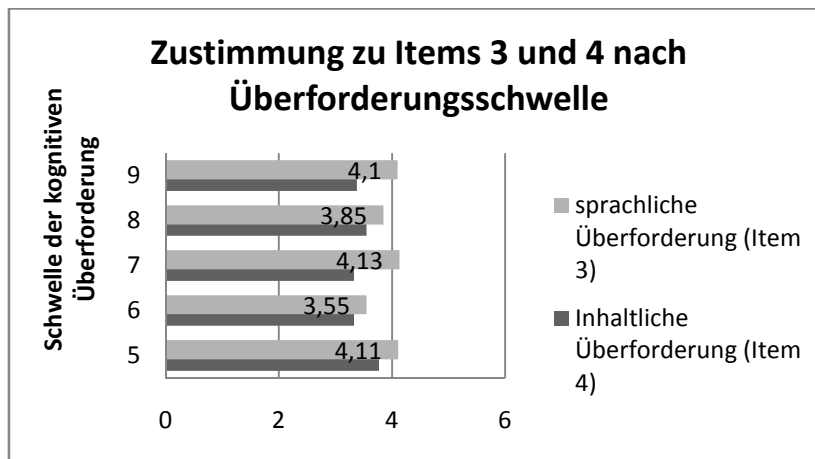


Abbildung 25: Zustimmung zu Items 3 und 4 geordnet nach Überforderungsschwelle (1= Stimmt genau; 5= Stimmt gar nicht; n=104)

Kompetenzwert weniger zugestimmt wird. Die gleiche Tendenz zeigt sich auch bei der inhaltlichen Überforderung. Tendenziell bereuen Schülerinnen und Schüler mit einem höheren Niveau von Bewertungskompetenz die Wahl von bilinguaem Biologieunterricht weniger als ihre Mitschüler mit einem geringeren Kompetenzwert. Interessanterweise bereut die Gruppe mit dem höchsten Kompetenzwert die Wahl jedoch am meisten.

Tabelle 31: Zustimmung zu Items Motivation, Überforderung Sprache, Überforderung Inhalt und Bili-Wahl bereit, nach Kompetenzwertgruppen (1= Stimmt genau; 5= Stimmt gar nicht; n=104)

	Motivation	Überforderung Sprache	Überforderung Inhalt	Wahl bereit
9-13	3,16	3,41	2,83	3,92
14-18	2,91	3,71	3,12	4,08
19-23	2,75	4,05	3,67	4,21
24-28	2,92	4,24	3,83	4,4
29-33	2,33	4,33	3,66	3,66

Die Tendenzen, die bei der Grupperung nach den erhobenen Werten von Bewertungskompetenz entsprechen eher den Erwartungen, als dies bei der Analyse anhand der Kognitionswerte der Fall war (vgl. Tab. 31). Scheinbar können Schülerinnen und Schüler bessere Einschätzungen anhand ihrer Kompetenzen machen. Eine hohe Kompetenz führt demnach auch zu einer hohen Selbsteinschätzung, die durch die erhobenen Daten dieser Studie bestätigt wird.

6 Diskussion und Ausblick

6.1 Diskussion der Ergebnisse

6.1.1 Allgemeine Ergebnisse

Aufgrund der Testergebnisse stellte sich heraus, dass es einen messbaren Zusammenhang zwischen der kognitiven Leistungsfähigkeit auf der einen und der Bewertungskompetenz auf der anderen Seite gibt. Zwar fällt die Korrelation dieser beiden Merkmale und der dadurch aufgezeigte Zusammenhang mit $r = .223$ nur gering aus, allerdings ist er auch nicht zu leugnen. Trotzdem bleibt festzuhalten, dass sich der erwartete deutliche Zusammenhang nicht bestätigt hat.

Studien zur Ausprägung und zum Zuwachs sprachlicher Kompetenzen haben wiederholt darauf hingewiesen, dass Schülerinnen und Schüler in bilingualen Kursen insgesamt eine höhere sprachliche Kompetenz in Englisch haben, als Schüler die keinen bilingualen Sachfachunterricht besuchen (vgl. BOHN 2008, FEHLING 2009, KRECHEL 2008). Auch der Kompetenzzuwachs für die sprachlichen Kompetenzen ist für die bili-Schüler deutlich erhöht. Dieser Aspekt wurde in dieser Studie jedoch nicht untersucht. Über den Zuwachs an Bewertungskompetenz kann diese Studie ebenfalls keine Aussage treffen, weil sie nur eine einmalige Befragung der Schülerinnen und Schüler angewendet hat. Der Zuwachs ist aber nur durch ein Pre/Post-Test-Setting messbar. Hier besteht weiterhin die Notwendigkeit weitergehender Forschungsansätze, um die Leistungen und Kompetenzen von bili-Schülern und mono-Schülern vergleichen zu können.

Aussagen über die Bewertungskompetenz als Ganzes sind ebenfalls problematisch. Zwar wird für die Berechnung der Korrelation jeweils ein Zahlenwert angegeben, der für das Ausmaß der Bewertungskompetenz steht, welches ein bestimmter Schüler aufweist, allerdings trifft dieser Wert keine Aussage darüber, welche Teilaspekte der Bewertungskompetenz ein Schüler oder eine Schülerin beherrscht. Daher wurde auf das Teilkompetenzmodell von REITSCHERT ET AL (2007) zurückgegriffen. Wie bereits in Kapitel 5 erwähnt, werden einzelne Teilkompetenzen in den Teilaufgaben spezifisch abgefragt. Dadurch kann man erkennen, in welchem Bereich schon eine hohe Kompetenz vorliegt und bei welchen Teilkompetenzen noch Arbeitsbedarf besteht. Auffällig gute Ergebnisse wurden bei den Teilkompetenzen „Argumentieren“, „Perspektivwechsel“ und „Urteilen“ erreicht. Zwei dieser drei

Teilkompetenzen sind keine rein auf den Biologieunterricht bezogenen Kompetenzen, weshalb sie auch in anderen Fächern regelmäßig eingeübt werden. Dieser Umstand trägt sicherlich dazu bei, dass die Schülerinnen und Schüler gerade bei diesen Teilkompetenzen die besten Ergebnisse erzielen können. Auch die Teilkompetenz „Urteilen“ zeichnet sich durch hohe Werte bei den Schülerleistungen aus. Davon sind die Teilkompetenzen „Argumentieren“ und „Bewusstmachen der eigenen Einstellung“ nicht zu trennen. Ein Teilnehmer kann nur dann ein hohes Niveau von Urteilskompetenz aufweisen, wenn er sich über seine eigene Einstellung bewusst wird und sein Urteil mit stichhaltigen Argumenten unterstützt.

Durchschnittliche Ergebnisse konnten bei den Teilkompetenzen „Beurteilen“ und „Folgenreflexion“ erzielt werden, wobei diese beiden Teilkompetenzen nicht immer trennscharf zu unterscheiden sind. Bei der Beurteilung von Situationen erzielten die Teilnehmer mittlere bis gute Ergebnisse während bei der Folgenreflexion ausschließlich mittlere Ergebnisse erreicht wurden. Das Problem der Schülerinnen und Schüler liegt bei dieser Teilkompetenz darin, dass hauptsächlich Folgen reflektiert werden, die kurzfristig auf unmittelbar betroffene Personen wirken. Eine Abstraktion auf gesellschaftliche Folgen, oder Folgen, die langfristig und auch auf mittelbar betroffene Personen wirken, findet nur selten statt. An dieser Stelle besteht Übungsbedarf, wie im folgenden Kapitel noch erläutert wird. Die Fähigkeit, abstrakte Folgen abzuwägen und zu vergleichen wird gerade im Hinblick der gesellschaftlichen Forderungen an den Biologieunterricht bedeutsam. Zukünftige Entscheidungsträger in Politik, Gesellschaft und Wissenschaft, zu denen Schülerinnen und Schüler im deutschen Bildungssystem geformt werden sollen, benötigen diese Kompetenz, um Vor- und Nachteile von Entscheidungsoptionen abzuwägen und potentielle Gefahren und Risiken zeitnah zu erkennen. REITSCHERT und HÖBLE benennen Voraussetzungen, um direkte und abstrakte Folgen abwägen zu können. „Die Nennung der klar abschätzbaren, unmittelbaren Folgen [...] korreliert eng mit dem nötigen Fachwissen und der Beurteilungsfähigkeit. Mögliche langfristige Folgen zu antizipieren erfordert Imaginationsvermögen und die Einnahme einer gesellschaftlichen Perspektive“ (2006, S. 104). Die Aufgabe 5B liefert ein gutes Beispiel: Natürlich birgt die Legalisierung von Stammzellenforschung gewisse Vorteile was den medizinischen und wissenschaftlichen Fortschritt angeht, allerdings dürfen auch gesellschaftliche

Gefahren nicht außer Acht gelassen werden. Beispielsweise bestehen Gefahren wie Missbrauch, sowie die Diskriminierung Kranker und Behinderter, falls deren Leiden für die Zukunft ausgemerzt würde. Dieses gründliche Abwägen zwischen persönlichen und gesellschaftlichen, Primär- und Sekundärfolgen, sowie eine Verbindung all dieser Aspekte in einer begründeten Entscheidung ist der Zielzustand, den Schülerinnen und Schüler nach Abschluss ihrer Schulbildung erreichen haben sollen.

Bei den beiden Teilkompetenzen „Wahrnehmen und Bewusstmachen moralischer Relevanz“ und „Ethisches Basiswissen“ erzielten die Schülerinnen und Schüler die schwächsten Ergebnisse. Viele Schüler waren nicht in der Lage, eine Dilemma-Situation hinsichtlich ihres ethisch-moralischen Wertkonflikts zu beschreiben. Der Hauptteil der Beschreibungen war rein pragmatisch-deskriptiv. Das bedeutet, dass der Mehrheit der Schüler diese Teilkompetenz fehlt oder unzureichend aufgebaut ist. Viele Methoden und Modelle zur Anwendung von Bioethik oder Nachhaltigkeit im Biologieunterricht hängen jedoch stark von dieser Teilkompetenz ab (vgl. BÖGEHOLZ, HÖBLE, LANGLET, SANDER & SCHLÜTER 2004). Die im Testheft verwendete Sechs-Schritt-Methode von CORINNA HÖBLE ist nur eine dieser Techniken, die ohne die Fähigkeit, eine moralisch-ethische Dimension zu erkennen, nicht angewendet werden kann. Ähnliches gilt für die Teilkompetenz „Ethisches Basiswissen“. Ohne diese Basis ist das Bewusstmachen und Wahrnehmen ethischer Relevanz stark erschwert. REITSCHERT und HÖBLE definieren den Inhalt dieser Teilkompetenz in der Fähigkeit, eine Situation begründet als moralisch relevante oder als reine Entscheidungssituation zu erkennen (vgl. 2006). Es ist auffällig, dass vor allem die Teilkompetenzen der Bewertungskompetenz schwach ausgebildet sind, die ihren Ursprung im Ethikunterricht haben. Dies ist sicherlich eine Folge von unzureichender Verknüpfung von Ethik- und Religionsunterricht mit dem Biologieunterricht. Ein gewisses ethisches Basiswissen kann und soll im Biologieunterricht vermittelt werden, allerdings müssen tiefergehende Betrachtungen auf den Religions- oder Ethikunterricht verschoben werden. Die Behandlung bioethischer Aspekte in diesen Fächern böte dazu sicherlich ausreichend Gelegenheit. Durch eine hohe Alltags- und Gesellschaftsrelevanz bioethischer Themen wären auch ein hohes Schülerinteresse und eine hohe Motivation zu erwarten.

6.1.2 Überprüfung der Hypothesen

Die erste Hypothese, die im Vorfeld formuliert wurde, bezieht sich auf den Vorteil älterer Schülerinnen und Schüler bei der Bearbeitung des Fragebogens. Dieser Vorteil wurde mit einer längeren Exposition zu biologischem und bioethischem Wissen und dem damit verbundenen Plus an Möglichkeiten der Schemagenerierung und –automatisierung begründet. Anhand der Ergebnisse dieser Studie kann diese Hypothese nur teilweise verifiziert werden. Zwar zeigen die 17-18-jährigen Teilnehmer bessere Ergebnisse bei den Testaufgaben als die 15-16-Jährigen, allerdings sind die Ergebnisse der 19-21-Jährigen nicht besser als die der 17-18-Jährigen. In einigen Fällen sind sie sogar schlechter als die der jüngsten Gruppe. Bei der kognitiven Leistungsfähigkeit sieht das Verhältnis identisch aus; den besten Wert hat die Gruppe der 17-18-Jährigen, gefolgt von der jüngeren Gruppe G1. Die Gruppe der ältesten Schülerinnen und Schüler hat den geringsten Kognitionswert. Die Werte unterscheiden sich jedoch nur geringfügig voneinander, sodass davon ausgegangen werden kann, dass die Stichprobengröße und andere statistische Faktoren einen hohen Einfluss auf die Werte haben.

Die prozentualen Anteile der nach Alter geteilten Stichprobe (vgl. Tab. 5-8) belegen die Durchschnittswerte. Wie bereits festgestellt, sind die prozentualen Anteile der jeweiligen Altersgruppe bei den Teilaufgaben gut vergleichbar. In neun von 15 Teilaufgaben haben alle drei Gruppen ihre Maxima auf dem gleichen Niveau und die Unterschiede in den Werten sind Resultat der geringen Stichprobengrößen der Gruppen G1 (n=23) und G3 (n=10). Einzig die Gruppe der 17-18-Jährigen kann in drei der Teilaufgaben deutlich besser abschneiden, als die anderen beiden Gruppen.

Auch die errechneten Korrelationswerte, die sich im Bereich einer schwachen Korrelation befinden, bestätigen diese Befunde. Die kognitive Leistungsfähigkeit hat nur in der ältesten Gruppe einen statistisch signifikanten Wert von $r = .364$. Dass dieser Wert im unteren Bereich der mittleren Korrelation angesiedelt ist, muss nicht weiter diskutiert werden. Es ist allerdings interessant, dass diese Korrelationsbefunde den Erwartungen widersprechen. Laut *Cognitive Load Theory* wäre eine hohe Korrelation bei den jungen Schülern zu erwarten gewesen, während bei der ältesten Gruppe die geringste Korrelation auftreten dürfte, weil diese Schülerinnen und Schüler faktisch mehr Schulzeit darauf verwenden konnten, erforderliche Schemata zu generieren und automatisieren bzw. erforderliche Techniken zu erlernen. Das

Anlegen von Vorwissen verringert laut *Cognitive Load Theory* den limitierenden Einfluss des Arbeitsgedächtnisses. Die logische Konsequenz wäre also, dass ältere Gruppen länger Zeit hatten, mehr und bessere Schemata generierten und damit die Limitierungen durch das Arbeitsgedächtnis ausschalten. Erschwerend kommt noch hinzu, dass die CLT vorrangig auf Novizen anwendbar ist, d.h. auch Schülerinnen und Schüler, die einen Unterrichtsinhalt zum ersten Mal bearbeiten und daher kein Vorwissen haben können. Für die Aufgaben des Testhefts ist diese Annahme künstlich und unrealistisch. Auch wenn die Aufgaben und Themen nicht im Biologieunterricht behandelt wurden, sind es dennoch Themen des öffentlichen bzw. privaten Interesses, von denen die meisten Schüler zumindest schon einmal gehört haben müssten. Gerade die älteren Schüler hätten eigentlich bereits mit den Themen in Kontakt kommen müssen.

Möglicherweise beinhaltet die Gruppe G3 Schülerinnen und Schüler, die ein Schuljahr wiederholen mussten und generell schlechtere Schulleistungen erbringen. Dies würde das schlechtere Abschneiden dieser Gruppe erklären. Tatsächlich hat G3 die geringste durchschnittliche Biologienote; auch die Mathematiknote ist durchschnittlich die geringste. Allerdings fallen die Werte der Gruppe G1 sehr gering aus (Bio: 8,08 (G1) zu 8,2 (G3); Mathe: 7,8 (G1) zu 7,1 (G3), in Biologie ist die Gruppe G3 sogar leicht besser als G1.

Um den Einfluss des Alters auf die Bewertungskompetenz genauer zu untersuchen, wurde die Gesamtstichprobe in G8- und G9-Schüler unterteilt. Bei diesen beiden Schülergruppen kann man mit Bestimmtheit sagen, dass die G8-Schüler ein Schuljahr weniger Zeit zur Generierung der Schemata hatten als ihre G9-Mitschüler. Die Ergebnisse dieser Gruppeneinteilung liegen mit wenigen Ausnahmen noch näher beieinander als dies schon bei der Einteilung nach Alter der Fall war. Wie im Ergebnisteil formuliert, liegt der Fall vor, dass mehr G9-Schüler auf Niveau III sind, als dies für G8-Schüler der Fall ist, allerdings gibt es auch Teilaufgaben bei denen dieses Verhältnis umgedreht ist. Insgesamt sind die Unterschiede nicht statistisch signifikant, was sich auch in den Korrelationswerten für beide Gruppen widerspiegelt. Die Korrelationen von $r = .231$ (G8-Gruppe) und $r = .218$ (G9-Gruppe) spiegeln beide einen geringen Zusammenhang von kognitiver Leistungsfähigkeit und Ausprägung der Bewertungskompetenz für die jeweilige Gruppe wieder. Allerdings entspricht der etwas geringere Wert der G9-Gruppe den

Erwartungen, die an eine Schülergruppe gestellt wird, die schon länger in der Schule ist. Der Unterschied ist allerdings zu gering um die Hypothese nachhaltig zu stützen. Ein hoher Einfluss sowohl des absoluten Alters der Schülerinnen und Schüler als auch des Bildungsganges und der damit bereits absolvierten Schulzeit auf die Bewertungskompetenz konnte durch die Studie nicht nachgewiesen werden. Weder bei einer willkürlichen Alterseinteilung, noch bei einer Einteilung nach Bildungsgang treten signifikante Leistungsunterschiede hervor. Die erste Hypothese ist somit widerlegt. Diese Beobachtung ist unabhängig von der Unterrichtssprache des Biologieunterrichts gültig.

Die zweite Hypothese besagt, dass die kognitive Mehrbelastung, der die Schülerinnen und Schüler in bilinguaem Biologieunterricht ausgesetzt sind, sich negativ auf die Leistungen und die Bewertungskompetenz dieser Schüler auswirkt. Grundlage dieser Hypothese ist der Umstand, dass die Schülerinnen und Schüler neben der kognitiven Belastung durch die biologischen und ethischen Fachinhalte zusätzlich noch einer sprachlichen Belastung ausgesetzt sind, die im Falle von Aufgaben, deren Bearbeitung und Lösung eine hohe Bewertungskompetenz erfordern zusätzlich erhöht wird. Die Bewertungskompetenz zeichnet sich durch die Verbindung von biologischen und bioethischen Aspekten, bzw. Aspekten der Nachhaltigkeit aus. Folglich entsteht auch eine kombinierte sprachliche Belastung. Laut *Cognitive Load Theory* hätten diese Belastungseinheiten zumindest im Falle der Schülerinnen und Schüler in bilingualen Kursen zu einer Überlastung führen müssen. Diese Überforderung würde sich in signifikant schlechteren Ergebnissen bei den Testheftausgaben widerspiegeln. Die Hypothese muss nach Auswertung der Ergebnisse weitgehend falsifiziert werden. Zwar tritt ein Leistungsunterschied zwischen den beiden Gruppen auf, aber dieser ist nicht groß genug, um von einer signifikant schlechteren Performanz der bili-Gruppe zu sprechen.

In den Ergebnissen wurde deutlich, dass beide Gruppen im ersten Teil des Fragebogens nahezu identische Ergebnisse abliefern (Aufgabe 1 bis 4B). In diesem Teil sind die bilingual unterrichteten Schüler sogar leicht besser als die der mono-Gruppe. Dieses Verhältnis dreht sich dann beim zweiten Teil (Aufgabe 5) wieder um. Hier erzielt die mono-Gruppe die eindeutig besseren Ergebnisse. Die Vorteile, die die bili-Gruppe im ersten Teil hat sind insgesamt kleiner als jene, die die mono-Gruppe im zweiten Teil hat. Somit kann man insgesamt von einer etwas besseren

Leistung derjenigen Schülerinnen und Schüler sprechen, die deutschsprachigen Biologieunterricht besuchen.

Auch die Korrelation der beiden Merkmale kognitive Leistungsfähigkeit und Bewertungskompetenz bildet diesen Trend ab. Mit $r = -.035$ besteht bei der mono-Gruppe eine sehr schwache negative Korrelation. Dieser Wert ist so gering, dass er nahezu keinen Zusammenhang zwischen den Merkmalen detektiert. Wäre er allerdings höher und weiterhin negativ, würde er der Annahme widersprechen, dass ein Zusammenhang nach dem Schema „Je höher die kognitive Leistungsfähigkeit ist, desto höher ist die Bewertungskompetenz ausgeprägt“ besteht. Der ausgesprochen geringe Wert reicht aber nicht aus, um einen je-höher-desto-niedriger-Zusammenhang zu konstruieren. Zwar liegt der Korrelationswert der bili-Gruppe mit $r = .349$ auch nur knapp oberhalb der Schwelle von schwacher bis mittlerer Korrelation, allerdings ist der Wert unabhängig von seiner isolierten Aussagefähigkeit zehnfach größer als der der mono-Gruppe. Das bedeutet dann, dass die kognitive Leistungsfähigkeit einen zehnfach höheren Einfluss auf die Bewertungskompetenz der bili-Schüler hat, als dies bei der mono-Gruppe der Fall ist. In diesem Aspekt muss die aufgestellte Hypothese also verifiziert werden.

Betrachtet man die Korrelationen der kognitiven Leistungsfähigkeit mit dem erzielten Ergebnis bei gewissen Teilkompetenzen, kann man weitere Schlüsse ziehen. Es wurde bereits diskutiert, dass einige Teilkompetenzen der Bewertungskompetenz besser ausgeprägt sind als andere. Vergleicht man nun die Ergebnisse, die die Schülerinnen und Schüler bei ausgewählten Teilaufgaben erreicht haben mit den Kognitionswerten, so werden zwei Aspekte deutlich.

Tabelle 32: Korrelationswerte für die Teilkompetenzen „Urteilen/Argumentieren“ und „Ethisches Basiswissen“ für die gesamte Stichprobe, sowie die bili-Gruppe und die mono-Gruppe

	1D + 5C	2 + 5A
Korrelation gesamt	$r = .175$	$r = .200$
Korrelation bili-Gruppe	$r = .291$	$r = .284$
Korrelation mono-Gruppe	$r = -.050$	$r = .015$

Verglichen werden zwei Teilkompetenzen, in denen die Teilnehmer stark abgeschnitten haben, nämlich „Urteilen/Argumentieren“ (1D+5C) und die Teilkompetenz, bei der die größten Mängel auftraten, „Ethisches Basiswissen“ (2+5A). Die Korrelation zwischen kognitiver Leistungsfähigkeit und

Teilbewertungskompetenz ist für die Teilkompetenz „Ethisches Basiswissen“ höher als für „Urteilen/Argumentieren“. Allerdings ist der Unterschied nur gering und beide Werte liegen im Bereich einer schwachen Korrelation. Viel interessanter ist aber der Vergleich zwischen der mono-Gruppe und der bili-Gruppe. Hier fallen die Unterschiede deutlich größer aus. Zwar liegen die Werte immer noch im Bereich einer schwachen Korrelation, allerdings liegen die Werte der bili-Gruppe in beiden Fällen knapp unterhalb der Schwelle zur mittleren Korrelation, während die Werte der mono-Gruppe eine sehr schwache bzw. keine Korrelation darstellt. Daraus kann man schließen, dass der Einfluss der kognitiven Leistungsfähigkeit auf die Ausprägung der Bewertungskompetenz innerhalb der einzelnen Teilkompetenzen schwankt und dass die kognitiven Limitierungen durch das Arbeitsgedächtnis eher Schülerinnen und Schüler der bili-Gruppe betreffen. Folglich trifft die *Cognitive Load Theory* in gewissem Maße hier zu, was ein weiteres Argument dafür liefert, dass die zweite Hypothese zumindest teilweise verifiziert werden muss. Im Hinblick auf die Testheftaufgaben bedeutet dies, dass viele Schülerinnen und Schüler von der zweiten Aufgaben kognitiv überfordert waren. Das bedeutet, dass ihr Arbeitsgedächtnis so mit dem Verstehen der Aufgabe beschäftigt war, dass für deren Lösung kaum bzw. keine Kapazität übrig blieb, während andere Aufgaben die Teilnehmer weniger stark kognitiv forderten und folglich erfolgreicher gelöst werden konnten.

Dadurch, dass die Korrelationswerte so gering ausfallen, kann man zwei mögliche Schlüsse ziehen. Einerseits kann es sein, dass die kognitive Leistungsfähigkeit zwar einen gewissen Einfluss auf die Bewertungskompetenz hat, dieser Einfluss aber geringer ist, als ursprünglich angenommen. Dieser eher generellen Annahme kann allerdings ein zweiter Schluss gegenübergestellt werden, nämlich, dass der Einfluss generell hoch ist, d.h. höher als in dieser Studie nachgewiesen. Das würde dann im Umkehrschluss bedeuten, dass der Einfluss kognitiver Charakteristika im Kontext dieser Studie einen geringeren Einfluss hatte, sei es, weil die Aufgaben keinen allzu hohen Anspruch stellten, oder weil die behandelten Inhalte doch kein thematisches Neuland mehr für die Teilnehmer darstellte. Wie bereits weiter oben erwähnt, gab es die meisten Schülerfragen zur Aufgabe 2 und den dort zu erstellenden Syllogismen. Die Nachfragen wurden sowohl von Teilnehmern der bili-Gruppe als auch von Teilnehmern der mono-Gruppe gestellt. Diese Technik der ethischen Logik war für

viele Schülerinnen und Schüler beider Untergruppen thematisch neu, weshalb die *Cognitive Load Theory* hier anwendbar ist. Bei bilingualen Schülerinnen und Schülern ist der Korrelationswert deutlich höher als bei den mono-Schülern, d.h. vor allem Schüler mit einem größeren Arbeitsgedächtnis konnten diese Aufgabe korrekt lösen, während Schülerinnen und Schüler mit einem kleineren Arbeitsgedächtnis eher dazu tendierten, die Aufgabe nur teilweise oder gar nicht korrekt zu bearbeiten. Diese Ergebnisse bedeuten nicht, dass die Aufgabenstellung für die bili-Gruppe schwerer, d.h. kognitiv anspruchsvoller, war als für die mono-Gruppe. Was ausgesagt wird ist, dass im Falle der bili-Gruppe die kognitive Mehrbelastung, die durch den bilingualen Biologieunterricht verursacht wird, noch zusätzlich zu der Belastung der Aufgabenstellung wirkt und dadurch zur höheren Korrelation führt.

Die Grundhypothese ging davon aus, dass bili-Schüler deutlich schlechtere Ergebnisse als mono-Schüler erzielen, wenn komplexe biologische Themen aus den Feldern Bioethik und Nachhaltige Entwicklung behandelt werden. Durch den Vergleich der Ergebnisse in den beiden Gruppen konnte diese Hypothese nicht verifiziert werden. Beide Gruppen zeigen vergleichbare Leistungen. Zwar hat die mono-Gruppe insgesamt leicht bessere Ergebnisse als die bili-Gruppe erzielt, allerdings war in einigen Teilaufgaben auch die bili-Gruppe etwas besser. Diese Ergebnisse decken sich auch mit den Befunden von KONDRING & EWIG (2005), bei deren Studie leicht bessere Leistungen der mono-Schüler detektiert wurden. Die Ergebnisse der Selbsteinschätzung der Überforderung der bili-Gruppe bestätigt, dass sich die Schülerinnen und Schüler, die bilingualen Biologieunterricht einer Überforderung durch die Unterrichtssprache und durch die Inhalte nicht zustimmen. Viele Schüler der bili-Gruppe würden erneut bilingualen Biologieunterricht belegen, was sich auch in einem hohen Motivationswert äußert (vgl. Kap. 5.3). Diese Ergebnisse sind deswegen bedeutsam, da Faktoren wie Motivation und Emotionen gegenüber dem Unterricht ebenfalls signifikante Einflüsse auf die Leistung von Schülerinnen und Schülern haben.

Durch die Korrelation der kognitiven Leistungsfähigkeit und der Bewertungskompetenz konnte für die bili-Gruppe ein zehnfach höherer Wert ermittelt werden. Die kognitive Leistungsfähigkeit hatte im Kontext des Testhefts jedoch nur einen geringen Einfluss auf die Leistungen, sprich die Bewertungskompetenz der Schülerinnen und Schüler, der höhere Korrelationswert

deutet jedoch darauf hin, dass bei noch komplexeren Unterrichtsinhalten dieser Einfluss an Gewicht gewinnt und die Performanz der bili-Schüler deutlichere Unterschiede zur mono-Gruppe aufweist.

Bei der Ergebnispräsentation wurde im Anschluss an den Vergleich der bili- und mono-Gruppe noch eine Verbindung mit dem Alter dargestellt. Die Faktoren Alter und Unterrichtssprache wurden kumuliert und in Tabelle 17-20 dargestellt. Zwar unterscheiden sich die Schülerleistungen auch hier wieder nicht signifikant, allerdings hat die kognitive Leistungsfähigkeit einen deutlich höheren Einfluss auf die Leistung. Bei der Kumulierung der Faktoren Alter und Unterrichtssprache wäre hinsichtlich der bislang präsentierten Ergebnisse zu erwarten, dass die Korrelation bei der Gruppe bG1 am höchsten, bei mG2 am niedrigsten sein müsste. Im Kontext des Fragebogens war zwar zu beobachten, dass mG2, wie auch mG1 eine sehr geringe Korrelation aufwiesen, allerdings hat die Gruppe bG2 mit $r = .460$ den höchsten Korrelationswert. Der Wert der Gruppe bG1 liegt mit $r = .268$ an der Schwelle zur mittleren Korrelation. Die Werte der bili-Gruppen sind deutlich höher als die der mono-Gruppen. Daraus lässt sich schließen, dass die Faktoren Alter und Unterrichtssprache zusammen einen höheren Einfluss auf die Bewertungskompetenz haben, auch wenn die Faktoren alleine gesehen zu keinen großen Unterschieden führen.

Die dritte Hypothese, dass Schülerinnen und Schüler eines bilingualen Zuges durch den vermehrten Kontakt zur englischen Sprache als Wissenschaftssprache und Transportmedium einer geringeren kognitiven Belastung ausgesetzt sind und folglich beim Fragebogen besser abschneiden kann ebenfalls nicht bestätigt werden. Bei den Ergebnissen fällt vielmehr auf, dass die Modul-Gruppe oftmals besser abschneidet als die Zug-Gruppe. Folglich müsste der Hypothese nicht nur widersprochen, sie müsste vielmehr umgedreht werden. Diese Ergebnisse entsprechen absolut nicht den Erwartungen. Man muss folglich davon ausgehen, dass die sprachlichen Anforderungen, die durch Aufgaben zur Bewertungskompetenz an die Schülerinnen und Schüler gestellt werden, auch dann keinen allzu großen Einfluss auf die Leistungen haben, wenn es sich dabei nicht nur um biologische, sondern auch um bioethische Fachsprache handelt. Dadurch, dass die Modul-Gruppe gleich gute bzw. bessere Ergebnisse bei den Aufgaben erzielen konnte scheint bewiesen zu sein, dass der herkömmliche Englischunterricht ausreicht, um auch auf komplexen bilingualen

Biologieunterricht vorzubereiten. Dies widerspricht der Studie von Zydati (2002), in welcher der Autor zu dem Befund kommt, dass konventioneller Englischunterricht nicht ausreicht, um Schlerinnen und Schlern die bentigten CALPs zu vermitteln, um am Diskurs im Sachfachunterricht teilzunehmen. Ein groer Mehrwert, ausgehend von einem regelmigen Sprachkontakt, wie er in bilingualen Zgen vorherrscht, kann im Rahmen dieser Studie nicht erkannt werden. Das soll allerdings nicht bedeuten, dass dieser nicht existiert, sondern nur, dass der Mehrwert in diesem Kontext einen vernachlssigbaren Einfluss auf die Schlerleistung hat.

Zustzlich sollte man in Erinnerung rufen, dass ein Hauptteil der Zug-Gruppe die 11. Klasse besucht, whrend alle Teilnehmer der Modul-Gruppe die 12. Klasse besucht. Zwar hat das Alter, wie oben dargestellt, einen geringeren Einfluss auf die Bewertungskompetenz als angenommen, allerdings zeigt schon die Kombination von Alter und Unterrichtssprache, dass zwei weniger bedeutsame Kriterien ihre Wirkungen addieren knnen. Mglicherweise liegt hier ein Erklrungsansatz fr das schlechtere Abschneiden der Zug-Gruppe.

Auf Grundlage der Ergebnisse dieser Arbeit kann festgestellt werden, dass in englischsprachigen Kursen die Zielsetzungen fr erfolgreichen Biologieunterricht angemessen erreicht werden knnen. Zwar hat die kognitive Leistungsfhigkeit von Schlerinnen und Schlern einen gewissen Einfluss auf die Bewertungskompetenz, allerdings fllt diese Korrelation nur schwach aus. Es ist zudem zu erwarten, dass kognitive Charakteristika von Schlern auch auf andere Kompetenzbereiche und Unterrichtsaspekte einen Einfluss hat. Die Ergebnisse zeigen, dass bilingual unterrichtete Schlerinnen und Schler nur unwesentlich schlechter abschneiden als ihre muttersprachlichen Mitschler. Fr die Cognitive Load bedeutet dies, dass sie entweder keiner erheblich hheren kognitiven Belastung ausgesetzt sind als ihre muttersprachlichen Mitschler, oder dass diese Belastung einen geringeren Einfluss auf ihre Leistungen hat als vermutet. Folglich war bei der Bearbeitung der Testhefte der Einfluss des Arbeitsgedchtnisses ebenfalls geringer als vermutet. Zur abschliesenden Klrung, ob und in wie weit eine kognitive Mehrbelastung durch bioethische Themen vorliegt, sind weiterreichende Untersuchungen notwendig. Ein weiteres positives Ergebnis der Studie ist auch, dass die wissenschaftliche Arbeitsweise auch im bilingualen Unterricht verwendet werden kann, ohne eine kognitive berforderung der Schlerinnen und Schler riskieren zu mssen. Dieses

Ergebnis hat allerdings nur eingeschränkte Gültigkeit. Um abschließend zu klären, ob naturwissenschaftliches Arbeiten ohne Überforderung möglich ist, müssten ähnliche Studien für die beiden Kompetenzbereiche Erkenntnisgewinnung und Kommunikation durchgeführt werden.

Ein weiteres Ergebnis der Untersuchung ist aber auch, dass die Bewertungskompetenz insgesamt immer noch nicht allzu stark ausgebildet ist. Trotz der Bildungsreform als Reaktion auf den PISA-Schock sind die Bildungsstandards und Kompetenzen, zumindest was den Teilbereich Bewertungskompetenz angeht, noch nicht in den Schulen angekommen. Der maximal erreichbare Kompetenzwert liegt bei 38 Punkten. Der durchschnittliche Kompetenzwert liegt mit 20,56 bzw. entsprechend 54%, nur knapp oberhalb der Hälfte der zu erreichenden Punkte. Mit 59 Schülerinnen und Schülern liegen 35% der Stichprobe unterhalb des 50%-Wertes der Bewertungskompetenz und hätten, wäre das Testheft eine Klausur gewesen, weniger als 5 Punkte erzielt. Die 75%-Schwelle liegt bei 28,5 Punkten. Bezieht man die Schülerinnen und Schüler mit ein, die 28 Punkte erreicht haben, liegen 14 Teilnehmer über diesem Wert, was 8,3% der Stichprobe entspricht. Der höchste erreichte Wert mit 33 Punkten entspricht immerhin 86,8% der möglichen Punktzahl. Insgesamt ist die Spanne des Kompetenzwertes von 9 bis 33 Punkte sehr hoch. Diese große Streuung der Ergebnisse erinnert stark an Ergebnisse der PISA Studie, die schon im Jahr 2000 eine große Streuung der Leistungen in Naturwissenschaften als ein Problem des deutschen Bildungssystems identifiziert haben (vgl. DEUTSCHES PISA-KONSORTIUM 2001). Im Hinblick auf den KMK-Beschluss der nationalen Bildungsstandards für den Mittleren Schulabschluss im Jahr 2004 stellen sich bei diesen Ergebnissen Ernüchterung ein.

Im Vorfeld der Datenerhebung hat eine der Lehrerinnen darauf verwiesen, dass in den untersuchten Kursen nur während einer Doppelstunde auf bioethische Aspekte eingegangen worden sei. Eine andere Lehrerin merkte an, dass die im Testheft abgefragten Inhalte nicht in den Vorgaben zum Landesabitur vermerkt seien und folglich bislang auch nicht behandelt wurden. Diese beiden Bemerkungen, wenn sie auch möglicherweise als prophylaktische Rechtfertigung für ein potentiell schlechtes Abschneiden der Schülerinnen und Schüler getätigt wurden, weisen auf einen wichtigen Aspekt hin. Zwar sind die Bildungsstandards implementiert, allerdings weist die Umsetzung im Umfeld der Schule noch Mängel auf. Folglich lassen sich

aus den Befunden dieser Studie auch Handlungsoptionen für den Biologieunterricht ableiten.

6.2 Bedeutung der Ergebnisse für den Biologieunterricht

Obwohl die Hypothesen dieser Arbeit falsifiziert wurden, kann man aus den Ergebnissen der Studie einige Schlüsse ziehen, die für den Biologieunterricht, sei es bilingual oder auf Deutsch, bedeutsam sein können.

Des Weiteren gibt diese Arbeit durch seine Anlage auch gewisse Informationen zum Erwerb von Bewertungskompetenz im Biologieunterricht. Auch zu diesem Thema stellt sich die Forschungslage noch sehr dürftig dar. Man kann also sagen, dass die Ergebnisse der vorliegenden Studie gewisse Handlungsanweisungen für den bilingualen Biologieunterricht, aber in gleicher Weise auch für den Biologieunterricht allgemein aufzeigen, wenn es darum geht, Bewertungskompetenz zu fördern.

6.2.1 Bedeutung der Ergebnisse für den bilingualen Biologieunterricht

Anhand dieser Studie wurde erwiesen, dass die Leistungen der Schülerinnen und Schüler, die bilingualen Biologieunterricht besuchen, vergleichbar mit ihren Mitschülern aus herkömmlichem Biologieunterricht sind, auch wenn sie sich etwas schwächer darstellen. Das Minimalziel, das der bilinguale Unterricht erreichen muss kann somit als erreicht angesehen werden. Dies kann und darf aber nicht das Ziel sein. Die hohen Erwartungen und Versprechungen von Mehrwert, die an bilingualen Unterricht gestellt werden, dürfen sich aus fachwissenschaftlicher Sicht nicht nur auf sprachliche Vorteile beschränken. Diese Vorteile wären nur eindimensional und könnten auch durch einige zusätzliche Stunden Englischunterricht erreicht werden, sofern während dieses Unterrichts auf das Erlernen und Anwenden der *Cognitive Academic Language Proficiency* (CALP) fokussiert wird. Auch ein biologisch-naturwissenschaftlicher Mehrwert muss generiert werden können.

Um einen Mehrwert generieren zu können muss zuerst abschließend sichergestellt werden, dass die Leistungen bilingualer Klassen und Kurse auf lange Sicht vergleichbar oder sogar besser sind als die Leistungen der Mitschüler. Doch wie ist das zu erreichen? Eine Faustformel kann man nicht aufstellen, aber es gibt gewisse Stellschrauben, an denen Lehrer und Schulen drehen können damit das Erreichen der

Minimalanforderung an den bilingualen Unterricht auf lange Sicht nicht zu Zitterpartie wird.

Organisatorisch und institutionell ändert sich die Stundentafel von Schülerinnen und Schülern, die bilinguale Züge belegen dahingehend, dass zu Beginn der Sekundarstufe eine zusätzliche Stunde Englisch und eine zusätzliche Stunde Sachfach erteilt wird. Man sollte die zusätzliche Englischstunde nicht darauf verwenden, regulären Englischunterricht zu erteilen. Der Fokus sollte präferentiell auf dem Erwerb von CALPs liegen, um sprachliche Hindernisse auf lange Sicht abzubauen. Dass es hierbei nicht nur um bioethisches Vokabular gehen soll, versteht sich von selber. Es sollte sich ebenfalls nicht um ein stupides Vokabellernen handeln, sondern immer in einen naturwissenschaftlichen Kontext eingebettet sein. Ein frühzeitiges Erlernen der englischen Fachterminologie für die Naturwissenschaften fördert nicht nur die wissenschaftliche Sprachkompetenz und baut sprachliche Hürden ab, sondern übt auch die wissenschaftliche Arbeitsweise ein. Zu diesem Zweck erscheint es auch als sinnvoll, dass die zusätzliche Stunde Englisch nicht von einem reinen Englischlehrer erteilt wird, sondern dass ein Lehrer, der bilingualen Unterricht erteilt, diese Aufgabe übernimmt.

Nun gibt es diese organisatorischen Anpassungen für Schulen mit bilingualen Modulen nicht oder nur kaum. Schülerinnen und Schüler, die sich entscheiden, ein Sachfach für einen gewissen Zeitraum bilingual zu belegen werden dementsprechend ins kalte Wasser geworfen. Sie müssen sich innerhalb kürzester Zeit auf die Verwendung der englischen Sprache als Medium einstellen. Dies ist gerade im Fall der Biologie problematisch. Wie in den theoretischen Überlegungen dargestellt, haben Schülerinnen und Schüler oftmals erst in der Oberstufe die Möglichkeit bilingualen Biologieunterricht zu wählen. Die thematischen Anforderungen der Biologie steigen mit der Oberstufe rapide an. Dieser Anstieg kann zusammen mit der sprachlichen Umstellung zu einer Überlastung führen. Daher sollten auch Schulen, die bilinguale Module anbieten, begleitend eine oder zwei weitere Englischstunden anbieten, in denen die sprachlichen Hürden abgebaut werden können. Eine solche Methode wäre sicherlich sinnvoller, als das Arbeitsgedächtnis der Schüler mit langen Glossaren zu belasten. Wie der *Split-Attention-Effect* besagt, können zu viele Hilfsangebote während des Unterrichts eher kontraproduktiv sein, weil die Schülerinnen und Schüler ihre Aufmerksamkeit zwischen den verschiedenen Quellen

aufteilen müssen. Diese Befunde richten sich aber nicht generell gegen Vokabellisten, die ein erprobtes Instrument im Englisch- sowie im bilingualen Unterricht darstellen. Auch in Verbindung mit dem *Expertise-Reversal-Effect* ist ein additiver Englischunterricht in Form eines Propädeutikums sinnvoller, da die Schülerinnen und Schüler frühzeitig Schemata zur Wissenschaftssprache anlegen und keine Vokabelhilfen mehr benötigen. Für diese Schüler wären die Listen eher hinderlich, da sie sie nicht bräuchten, sie aber dennoch konsultieren würden, um sicher zu gehen.

Diese Betrachtungen sind keineswegs auf sprachliche Aspekte beschränkt. Vielmehr sollte auch die zusätzliche Sachfachstunde zur Einübung fachspezifischer Kompetenzen verwendet werden, um im Verlauf der Schulkarriere möglichst viele Hürden überspringen zu können. Problematisch ist hierbei wieder, dass im Falle eines bilingualen Zuges dasjenige Sachfach eine Extrastunde zugeteilt bekommt, das zuerst bilingual unterrichtet wird. Im Falle des Faches Geographie können somit auch biologisch relevante Arbeitsweisen erlernt werden. Wird jedoch Geschichte als erstes Fach bilingual belegt, so treten keine Synergie-Effekte auf. Im Fall der modularen Organisation gibt es die Extrastunde im Sachfach erst gar nicht. Eine solche Extrastunde sollte für jedes bilingual unterrichtete Sachfach erteilt und ausdrücklich dafür verwendet werden, die fachliche Arbeitsweise kennenzulernen. Dabei muss diese Extrastunde nicht über die gesamte Belegdauer eines bilingualen Fachs bestehen bleiben. Die Erteilung einer zusätzlichen Englisch- sowie einer zusätzlichen Sachfach-Stunde im ersten Lernjahr eines bilingual belegten Faches sollten ausreichen, um gewisse Kompetenzen und Schemata auszubilden.

6.2.2 Bedeutung der Ergebnisse für den Erwerb der Bewertungskompetenz im Biologieunterricht

Eine Erkenntnis dieser Studie ist, dass bei der Implementierung und Förderung Bewertungskompetenz noch Optimierungsbedarf besteht. Das generelle Abschneiden der Schülerinnen und Schüler ist nicht zufriedenstellend, wie weiter oben bereits dargestellt wurde. Der Umstand, dass 35% der Teilnehmer weniger als die Hälfte der möglichen Kompetenzpunkte erreicht ist ein erschreckendes Ergebnis. Es bestätigt in gewissem Maße die Befunde der PISA-Studien, dass im deutschen Bildungssystem zu wenig auf die Anwendung und Prozessorientierung fokussiert wird. Das schwache generelle Abschneiden sagt vor allem aus, dass aus einem vorhandenen Fachwissen,

sei es in den Köpfen der Schüler oder in einer schriftlichen Quelle, wie im Fall des Testheftes, keine ausreichend begründeten und durchdachten Urteile gefällt werden können.

Wie im Ergebnisteil dargestellt, liegt das Problem jedoch nicht im Urteil oder im Argumentieren, sondern im ethischen Basiswissen, dem Wahrnehmen moralischer Relevanz und teilweise in der Folgenreflexion. Auf diese Teilkompetenzen muss folglich das Hauptaugenmerk liegen, um Bewertungskompetenz bei den Schülern zu verankern. Diese Befunde können mit den Aussagen von EGGERT & BÖGEHOLZ verglichen werden, die die bisherige Forschung dahingehend zusammenfassen, „dass Schüler(innen) Schwierigkeiten im Umgang mit einer systematischen Vorgehensweise in Entscheidungsprozessen haben und über kaum Metastrategie- bzw. Bewertungsstrukturwissen verfügen“ (2006:191). Die beiden Autoren fassen weitergehend zusammen, dass ein systematisches Vorgehen ohne das Vorhandensein aller Teilkompetenzen der Bewertungskompetenz nicht möglich sei und dass die getroffenen Entscheidungen folglich intuitiv und rechtfertigend seien (vgl. *ibid.*).

Für die Förderung der beiden Teilkompetenzen „Ethisches Basiswissen“ und „Wahrnehmen und Bewusstmachen moralischer Relevanz“ bietet sich vor allem eine engere Zusammenarbeit mit den Fächern Religion und Ethik an. In diesen Fächern, vor allem in Ethik, sollten vermehrt bioethische Themen behandelt werden. Bei der Teilkompetenz „Folgenreflexion“ liegt das Problem nicht darin, dass keine Folgen reflektiert werden, sondern dass entweder ein zu geringer Abstraktionsgrad vorliegt oder Bereiche von Folgen nicht bedacht werden. Das Resultat der ersten Situation sind hauptsächlich persönliche, individuelle Folgen für wenige, direkt betroffene Personen, in der zweiten Situation werden entweder die individuellen oder die gesellschaftlichen Folgen vergessen. Folgenreflexion heißt jedoch, dass beide Bereiche bedacht werden. Hier sollte in Unterrichtssituation gezielt geübt werden einen steigenden Abstraktionsgrad zu erreichen und für die Folgenreflexion verschiedene Perspektiven einzunehmen, um eine möglichst umfassende Reflexion durchzuführen.

6.3 Methodenkritik

Natürlich kann diese Arbeit mit seiner Fragestellung keine abschließende Aussage über den Einfluss von kognitiver Leistungsfähigkeit auf die Leistungen von

Schülerinnen und Schülern im bilingualen Biologieunterricht treffen. Es war vielmehr Ziel der Arbeit, einen Einblick darüber zu geben, welcher Forschungsbedarf in der Fachliteratur besteht und wie diese Lücken gefüllt werden können. Auch wenn die vorliegenden Ergebnisse besagen, dass zwischen der kognitiven Leistungsfähigkeit und der Bewertungskompetenz von Schülerinnen und Schülern nur ein schwacher Zusammenhang besteht, ist nicht auszuschließen, dass kognitive Charakteristika einen maßgeblichen Einfluss auf den Lernerfolg und den Kompetenzerwerb im Fach Biologie haben. Schließt man in diese Einschränkungen noch mit ein, dass nur ein Kompetenzbereich untersucht wurde, und dies anhand einer relativ kleinen Stichprobe, wird schnell deutlich, dass eine abschließende Beantwortung der Frage nach einem Zusammenhang zwischen kognitiver Leistungsfähigkeit und den Kompetenzen der Schülerinnen und Schüler hier nicht abschließend möglich ist.

Während der Vorbereitung und Durchführung der Schülertests sind einige Mängel und Probleme aufgetaucht, die hätten verhindert werden können. Ein Pre-Test des gesamten Testhefts, und nicht nur des Kognitionstests wäre sinnvoll gewesen und hätte bereits im Vorfeld dabei geholfen, einige der Fehler aufzuheben, was möglicherweise zu aussagekräftigeren Ergebnissen geführt hätte. Ein Hauptproblem des Kognitionstests wurde bereits im entsprechenden Kapitel der Ergebnispräsentation angesprochen, nämlich die genaue Determinierung der Kognitionsschwelle. Eine andere Methode, um diese festzulegen hätte zweifelsfrei auch verwendet werden können, allerdings musste für die Auswertung eine Entscheidung getroffen werden. Inwieweit die ermittelte Kognitionsschwelle wirklich der kognitiven Leistungsfähigkeit der Schülerinnen und Schülern entspricht ist natürlich fragwürdig, weil die erhobenen Daten immer nur eine Momentaufnahme sein können. Um verlässliche Daten zu generieren, müsste der Kognitionstest mehrere Male durchgeführt werden und zwar jeweils mit anderen, zufälligen Zahlenfolgen. Dennoch stellt der verwendete Kognitionstest im Rahmen dieser Arbeit die geeignetste Methode dar.

Während der Datenerhebung in den Schulklassen wurde es versäumt, spezifische Aufzeichnungen über die Schülerfragen sprachlicher und inhaltlicher Natur zu machen. Diese Aufzeichnungen hätten bei der Auswertung der Ergebnisse sinnvoll sein können. Besonders viele sprachliche Nachfragen in einem bilingualen Kurs

hätten einen interessanten Vergleichswert zu der Selbsteinschätzung sein können, um nur ein Beispiel zu nennen.

Eine Beobachtung, die allerdings während der Datenerhebung immer wieder gemacht wurde, sind Nachfragen zu den Aufgaben 2 und 3. Besonders bei Aufgabe 2 war vielen Schülern der Arbeitsauftrag unklar. Des Weiteren war die Methode des Syllogismus weitgehend unbekannt. Eine ausführlichere Beschreibung in der Aufgabenstellung hätte dieses Problem leicht beheben können. Da dies nicht der Fall war, könnten die Ergebnisse verfälscht sein. Schülerinnen und Schüler, die nachgefragt haben, hatten einen Vorteil gegenüber denen, die nicht nachgefragt haben. Zwar wurde darauf geachtet, bei Nachfragen nicht zu viel zu verraten, um ein gewisses Maß an Chancengleichheit aller Teilnehmer zu gewährleisten, allerdings besteht hier zweifellos ein Ungleichgewicht. Über das Ausmaß dieses Ungleichgewichts hätten die oben erwähnten, aber versäumten Aufzeichnungen Aufschluss geben können. Außerdem hätte in der Aufgabenstellung vermerkt werden sollen, dass jede der acht Aussagen nur einmal verwendet werden soll. Dieses Versäumnis trug sicherlich mit dazu bei, dass zwar viele Schüler den Syllogismus aufstellen konnten, der eine Eizellspende fordert, während der andere nur selten aufgestellt werden konnte.

Im Vorfeld der Datenerhebung stand der Prozess der Stichprobenakquise. Mit der letztendlichen Stichprobe kann man sicherlich sehr zufrieden sein, allerdings wäre diese in einem offizielleren Rahmen, beispielsweise im Rahmen einer Doktorarbeit oder eines universitären Forschungsansatzes sicherlich größer ausgefallen. Die meisten Lehrpersonen der Biologiekurse erhoben den Anspruch, den Fragebogen schon lange vor der Durchführung zu Anschauungszwecken ausgehändigt zu bekommen. Dieses Verhalten ist zwar verständlich, birgt aber gleichzeitig auch die Gefahr einer Verfälschung der Ergebnisse. Zwar weisen diese auf den ersten Blick nicht darauf hin, dass bestimmte Klassen einige der Aufgaben schon vorher in der abgefragten, oder einer ähnlichen Form eingeübt hatten, allerdings ist diese Möglichkeit auch nicht auszuschließen.

Zwar wurde die Untersuchung an vier Schulen durchgeführt, was zu einer Stichprobengröße von $n=169$ führte, allerdings reicht diese Stichprobe nicht aus, um wirklich aussagekräftige Daten erheben zu können. Das Hauptproblem für die

Auswertung bestand darin, dass sich die Stichprobe zwar gut in Untergruppen einteilen ließ, die dann eine vergleichende Betrachtung der Ergebnisse zuließ, jedoch zu dem Preis, dass diese Gruppen stark ungleiche Größen aufwiesen. Schon der Vergleich von bilingual unterrichteten Schülern (n=105) zu Schülerinnen und Schülern, die Biologie auf Deutsch belegen (n=64) ist unausgeglichen. Dieses Ungleichgewicht wird bei den anderen beiden Unterfragestellungen noch deutlicher ausgefallen. Die sehr geringen Stichprobengrößen bei den Altersgruppen der 15-16-Jährigen, sowie den 19-21-Jährigen lassen nur wenig aussagekräftige Ergebnisse zu. Eine bessere Auswahl der Stichprobe hinsichtlich einer gleichmäßigeren Verteilung auf Gruppen für die Unterfragestellungen wäre sinnvoll gewesen.

7 Literaturverzeichnis

BADDELEY, A. (1992). Working memory. *Science*, (255), 556–559.

BÖGEHOLZ, S.; HÖBLE, C.; LANGLET, J.; SANDER, E. & SCHLÜTER, K. (2004). Bewerten - Urteilen - Entscheiden im biologischen Kontext: Modelle in der Biologiedidaktik. *Zeitschrift für Didaktik der Naturwissenschaften. Biologie, Chemie, Physik*, 10, 89–115.

BOHN, M. (2008). Didaktische, methodische und organisatorische Herausforderungen auf dem Weg zum bilingualen Biologieunterricht der Zukunft. In SCHEERSOI, A. & KLEIN, H.P. (Hrsg.), *Bilingualer Biologieunterricht. Didaktik der Biowissenschaften. Frankfurter Beiträge zur biologischen Bildung* 6 (pp. 25–47). Aachen: Shaker.

BRÄKLING, A. (2007). Memory Capacity and Cognitive Load Test. Aus dem Internet: http://www.braekling.de/testlab/memory_capacity_cognitive_load_test.html. Abgerufen am 05.03.2012.

BRÜNKEN, R.; PLASS, J. L. & LEUTNER, D. (2003). Direct Measurement of Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 53–61.

CHANDLER, P. & SWELLER, J. (1992). The Split-Attention Effect As A Factor In The Design Of Instruction. *British Journal of Educational Psychology*, 62, 233–246.

CUMMINS, J. (1980). The crosslingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL quarterly*, 14(2), 175–187.

DESI-KONSORTIUM (Hrsg.) (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch*. Weinheim: Beltz PVU Verlag.

DEUTSCHES PISA-KONSORTIUM (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.

EGGERT, S. & BÖGEHOLZ, S. (2006). Göttinger Modell der Bewertungskompetenz - Teilkompetenz "Bewerten, Entscheiden und Reflektieren" für Gestaltungsaufgaben Nachhaltiger Entwicklung. *Zeitschrift für Didaktik der Naturwissenschaften. Biologie, Chemie, Physik*, 12, 177–197.

EGGERT, S. & HÖBLE C. (2006). Bewertungskompetenz im Biologieunterricht - Ein Überblick. *Praxis der Naturwissenschaften - Biologie in der Schule*, 55(1), 1–9.

EGGERT, S.; BÖGEHOLZ S.; WATERMANN, R. & HASSELHORN M. (2010). Förderung von Bewertungskompetenz im Biologieunterricht durch zusätzliche metakognitive Strukturierungshilfen beim Kooperativen Lernen - Ein Beispiel für Veränderungsmessung. *Zeitschrift für Didaktik der Naturwissenschaften. Biologie, Chemie, Physik*, 16, 299–314.

ELSTER, D. (2011). Linda's and Bob's Decision. *Unterricht Biologie*, (367/368), 35–44.

- FEHLING, S. (2008). *Language Awareness und bilingualer Unterricht: Eine komparative Studie*. Frankfurt am Main: Peter Lang GmbH.
- FEHLING, S. (2009). Lernprozesse und kognitive Entwicklung im bilingualen Unterricht: Bericht aus einer zweijährigen Longitudinalstudie. In D. CASPARI, W. HALLET, WEGNER, A. & ZYDATIB, W. (Hrsg.), *Bilingualer Unterricht macht Schule. Beiträge aus der Praxisforschung* (pp. 51–63). Frankfurt am Main: Peter Lang GmbH.
- FENKERT, A. (2009). *Bilingualer Sachfachunterricht: Curriculumsanspruch und Unterrichtswirklichkeit am Beispiel der Zweisprachigen Handelsakademie in Klagenfurt*. Saarbrücken: VDM Verlag Dr. Müller AG & Co. KG.
- HARMS, U.; MAYER, J.; HAMMANN, M.; BAYRHUBER, H. & KATTMANN, U. (2004). Kerncurriculum und Standards für den Biologieunterricht in der gymnasialen Oberstufe. In H.-E. TENORTH (Hrsg.), *Kerncurriculum Oberstufe II. Biologie, Chemie, Physik, Geschichte, Politik. Expertisen - im Auftrag der Ständigen Konferenzen der Kultusminister* (pp. 22–84). Weinheim: Beltz Verlag.
- HEMMELGARN, M. & EWIG, M. (2003). Bilingualer Biologieunterricht. Ein Forschungsfeld (auch) für die Biologiedidaktik. *Berichte des Institutes für Didaktik der Biologie der Westfälischen Wilhelms-Universität Münster, IDB*, (12), 39–62.
- HESSISCHES KULTUSMINISTERIUM (2010a). *Lehrplan Biologie. Gymnasialer Bildungsgang: Jahrgangsstufen 5G bis 9G und gymnasiale Oberstufe*. Aus dem Internet:
http://www.kultusministerium.hessen.de/irj/HKM_Internet?cid=ac9f301df54d1fbfab83dd3a6449af60.
- HESSISCHES KULTUSMINISTERIUM (2010b). *Lehrplan Englisch. Gymnasialer Bildungsgang: Jahrgangsstufen 5G bis 9G und gymnasiale Oberstufe*. Aus dem Internet:
http://www.kultusministerium.hessen.de/irj/HKM_Internet?cid=ac9f301df54d1fbfab83dd3a6449af60.
- HESSISCHES KULTUSMINISTERIUM (2012). *Zweisprachige Bildungsangebote an hessischen Schulen. Stand: Januar 2012*. Aus dem Internet:
http://www.kultusministerium.hessen.de/irj/HKM_Internet?cid=2937031b32ec515ee8037796d38bcac3, Zugriff am 23.03.2012.
- HÖBLE, C. & BÖGEHOLZ, H. (2006). Sechs Schritte moralischer Urteilsbildung: Aktuelle Beispiele aus der Bioethikdebatte. *Praxis der Naturwissenschaften - Biologie in der Schule*, 55(4), 1–6.
- HÖBLE, C. (2007). Theorien zur Entwicklung und Förderung moralischer Urteilsfähigkeit. In KRÜGER, D. & VOGT, H. (Hrsg.), *Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden* (pp. 197–208). Heidelberg: Springer Verlag.

KIRSCHNER, P. A.; SWELLER, J. & CLARK, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), 75–86.

KOCH, A. & BÜNDER, W. (2006). Fachbezogener Wissenserwerb im bilingualen naturwissenschaftlichen Unterricht. *Zeitschrift für Didaktik der Naturwissenschaften. Biologie, Chemie, Physik*, 12, 67–76.

KONDRING, B. & EWIG, M. (2005). Aspekte der Leistungsmessung im bilingualen Biologieunterricht. *Berichte des Institutes für Didaktik der Biologie der Westfälischen Wilhelms-Universität Münster, IDB*, 14, 49–62.

KRECHEL, H. L. (2008). Anmerkungen zur Didaktik des bilingualen Sachfachunterrichtes. In SCHEERSOI, A. & KLEIN, H. P. (Hrsg.), *Bilingualer Biologieunterricht. Didaktik der Biowissenschaften. Frankfurter Beiträge zur biologischen Bildung* 6 (pp. 1–24). Aachen: Shaker.

KULTUSMINISTERKONFERENZ (KMK) (2005). *Beschlüsse der Kultusministerkonferenz: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.

LEIBNIZ-INSTITUT FÜR DIE PÄDAGOGIK DER NATURWISSENSCHAFTEN (IPN) (2009). *bik - Biologie im Kontext: CD-ROM mit bik-Aufgaben*.

LIND, G. (2006). Das Dilemma liegt im Auge des Betrachters: Zur Behandlung bio-ethischer Fragen im Biologieunterricht mit der Konstanzer Methode der Dilemmadiskussion. *Praxis der Naturwissenschaften - Biologie in der Schule*, 55(1), 10–16.

MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.

MITTELSTEN SCHEID, N. & HÖBLE, C. (2007). Bewerten im Biologieunterricht: Niveaus von Bewertungskompetenz. Auf dem Weg vom normativen zum deskriptiven Kompetenzstrukturmodell. *Erkenntnisweg Biologiedidaktik*, 6, 87–104.

NOLD, G., HARTIG, J., HINZ, S. & ROSSA, H. (2008). Klassen mit bilingualem Sachfachunterricht: Englisch als Arbeitssprache. In DESI-KONSORTIUM (Hg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch*. Weinheim: Beltz PVU Verlag.

OSTERHAGE, S. (2009). Sachfachkönnen (scientific literacy) bilingual und monolingual unterrichteter Biologieschüler: ein Kompetenzvergleich. In D. CASPARI, W. HALLET, WEGNER, A. & ZYDATIß, W. (Hrsg.), *Bilingualer Unterricht macht Schule. Beiträge aus der Praxisforschung* (pp. 41–50). Frankfurt am Main: Peter Lang GmbH.

PAAS, F.; RENKL, A. & SWELLER, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38(1), 1–4.

PAAS, F.; RENKL, A. & SWELLER, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science*, 32, 1–8.

REITSCHERT, K. & HÖBLE, C. (2006). Die Struktur von Bewertungskompetenz: Ein Beitrag zur Dimensionierung eines Kompetenzmodells im Bereich der Bioethik. *Erkenntnisweg Biologiedidaktik*, 5, 99–114.

REITSCHERT, K., LANGLET, J., HÖBLE, C., & MITTELSTEN SCHEID, N. & SCHLÜTER, K. (2007). Dimensionen Ethischer Urteilskompetenz: - Dimensionierung und Niveaunkretisierung -. *Der mathematische und naturwissenschaftliche Unterricht*, 60(1), 43–51.

RENKL, A. & ATKINSON, R. K. (2003). Structuring the Transition From Example Study to Problem Solving in Cognitive Skill Aquisition: A Cognitive Load Perspective. *Educational Psychologist*, 38(1), 15–22.

RICHTER, R. & ZIMMERMANN, M. (2003). Und es geht doch: Naturwissenschaftlicher Unterricht auf Englisch. In M. OTTEN & E. WILDHAGE (Ed.), *Praxis des bilingualen Unterrichts* (pp. 116–146). Berlin: Cornelsen Scriptor.

SWELLER, J. (1994). Cognitive Load Theory, Learning Difficulty, and Instructional Design. *Learning and Instruction*, 4, 295–312.

SWELLER, J. (2005). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. MAYER (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 19–30). New York: Cambridge University Press.

SWELLER, J. (2006). How the Human Cognitive System Deals with Complexity. In ELEN, J. & CLARK, R. E. (Hrsg.), *Handling Complexity in Learning Environments: Theory and Research* (pp. 13–25). Oxford: Elsevier.

VAN MERRIËNBOER, J. J. G.; KIRSCHNER, P. A. & KESTER L. (2003). Taking the Load Off a Learner's Mind: Instructional Design for Complex Learning. *Educational Psychologist*, 38(1), 5–13.

VISSER, E. & HÖBLE, C. (2010). Bewerten bewerten: Diagnoseaufgaben für die Bewertungskompetenz im Biologieunterricht. *Der mathematische und naturwissenschaftliche Unterricht*, 63(5), 286–291.

WERNER, B. (2009). Entwicklungen und aktuelle Zahlen bilingualen Unterrichts in Deutschland. In D. CASPARI, W. HALLET, & WEGNER, A. & ZYDATIß, W. (Hrsg.), *Bilingualer Unterricht macht Schule. Beiträge aus der Praxisforschung* (pp. 19–28). Frankfurt am Main: Peter Lang GmbH.

YOUSOOF, M.; SAPIYAN, M. & KHAJA K. (2007). Measuring Cognitive Load - A Solution to Ease Learning of Programming. *World Academy of Science, Engineering and Technology*, 26, 216–219.

ZOHAR, A. & NEMET. F. (2002). Fostering Students' Knowledge and Argumentation Skills Through Dilemma in Human Genetics. *Journal of Research in Science Teaching*, 39(1), 35–62.

ZYDATIB, W. (2002). *Leistungsentwicklung und Sprachstandserhebungen im Englisch-unterricht. Methoden und Ergebnisse der Evaluierung eines Schulversuchs zur Begabten-förderung: Gymnasiale Regel- und Expressklassen im Vergleich.* Frankfurt am Main: Peter Lang GmbH.

Erklärung

Ich versichere hiermit, dass ich diese Arbeit selbstständig verfasst, keine anderen als die angegebenen Hilfsmittel verwendet und sämtliche Stellen, die benutzten Werken dem Wortlaut oder dem Sinne nach entnommen sind, mit Quellenangaben kenntlich gemacht habe. Dies gilt auch für Zeichnungen, Skizzen, Notenbeispiele und bildliche Darstellungen.

Kassel, den 20.06.2012

Daniel Busch

Anhang

1. Fragebogen Version A Deutsch
2. Fragebogen Version A Englisch
3. Erwartungshorizont mit Codierungsanleitung

Anhang 1: Fragebogen Version A Deutsch

TESTHEFT ZUR UNTERSUCHUNG EINER KORRELATION VON KOGNITIVER LEISTUNGSFÄHIGKEIT UND DER LÖSUNG VON BILINGUALEN BIOLOGIEAUFGABEN

Liebe Schülerin, lieber Schüler,

ich bitte dich, diesen Testbogen für meine Examensarbeit auszufüllen.

- Lies dir die Aufgaben genau durch
- Bitte versuche, die Aufgaben so gut wie möglich zu lösen. Wenn du bei einer Aufgabe oder einem Aufgabenteil nicht sicher bist, formuliere bitte trotzdem eine Antwort. In vielen Fällen gibt es keine richtigen oder falschen Antworten, sondern es kommt auf die Begründung an.
- Bitte achte darauf, dass du keine Aufgabe auslässt.
- Wenn dir etwas unklar ist, kannst du mich jederzeit ansprechen.

Deine Antworten werden von mir zwar ausgewertet, aber nicht bewertet. Das bedeutet dass dein Abschneiden bei den Aufgaben keinen Einfluss auf deine Biologienote haben wird. Deine Antworten werden natürlich auch von mir anonym behandelt.

Herzlichen Dank für deine Mitarbeit und viel Spaß!

Zuerst brauche ich einige Angaben von dir:

Alter: _____ Jahre

Geschlecht: männlich weiblich

Jahrgang: 10 11 12 13

Biologieunterricht: Biologie auf Deutsch Biologie bilingual

Seit wann belegst du bilingualen Biologieunterricht? _____

Belegst du auch noch andere bilinguale Fächer? nein ja: _____

Noten/Punkte im letzten Zeugnis: Biologie: _____ Englisch: _____
Deutsch: _____ Mathe: _____

TEIL 1: ERFASSUNG DER KONZENTRATIONSFÄHIGKEIT.

Dir wird mehrmals hintereinander eine kurze, willkürliche Abfolge von Zahlen gezeigt, **die du dir merken sollst, jedoch ohne sie aufzuschreiben**. Im Anschluss an jede Zahlenfolge werden dir einfache mathematische Rechnungen präsentiert, deren Lösungen du bitte unten notierst. Danach schreibst du bitte die Zahlenfolge auf.

Dieser Vorgang wird mehrmals wiederholt.

Bei diesem Test ist es wichtig, dass du nicht schummelst! Schreibe die Zahlenfolge bitte wirklich erst nach der Bearbeitung der Matheaufgaben auf.

Durchgang 1:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

Durchgang 2:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

Durchgang 3:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

Durchgang 4:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

TEIL 2: LEISTUNGSTEST BIOLOGIEAUFGABEN.

Aufgabe 1: Lindas und Bobs Entscheidung

Linda ist 16 Jahre alt und ungewollt in der neunten Woche schwanger. Sie und ihr Freund Bob, ein Austauschschüler aus Oklahoma, sind sich nicht sicher, ob sie das Baby bekommen sollen oder nicht. Sie wissen nicht genug über die Konsequenzen für die jeweiligen Entscheidungen. Lindas Eltern sind aus religiösen Gründen gegen eine Abtreibung. Sie möchten, dass ihre Tochter dem Kind das Leben schenkt und dass sie die Schule abschließt. Bob möchte Linda unterstützen, hat allerdings Angst vor der Verantwortung für die junge Familie. Er hat kein Einkommen und hofft in zwei Jahren die Schule abzuschließen. Bobs Eltern sind Geringverdiener. Linda und Bob fragen ihre Freunde um Hilfe bei der Entscheidung.

A) Definiere das vorliegende Dilemma

B) Nenne potentielle Lösungen des Dilemmas.

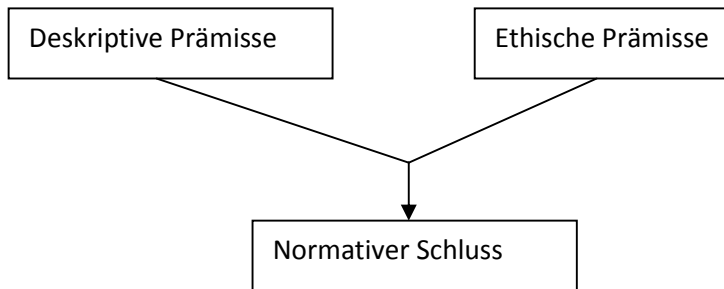
C) Nenne die ethischen Werte, die die jeweiligen potentiellen Lösungen betreffen.

D) Wie würdest du dich entscheiden? Gib Gründe!

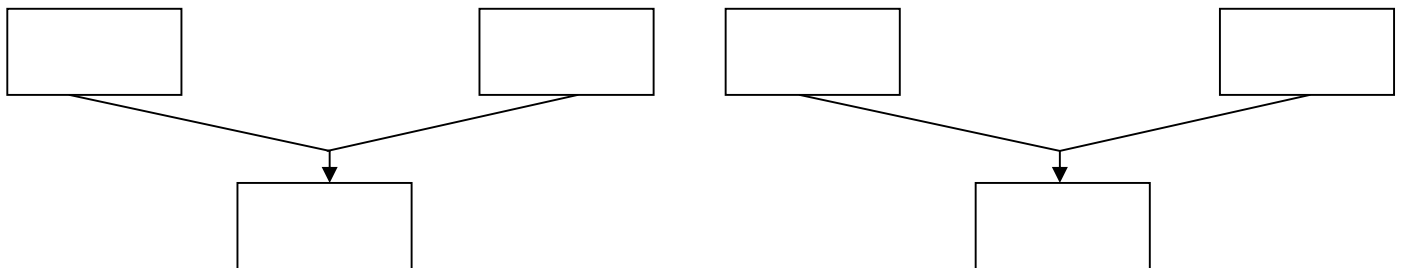
E) Liste die Konsequenzen deiner Entscheidung auf.

Aufgabe 2: Eizellspende

A) Konstruiere zwei logische Argumente aus den unten genannten Argumenten. Folge dabei diesem Schema.



A	Eizellspende erlaubt es einigen Paaren, ihren Kinderwunsch zu erfüllen.
B	Eizellspende sollte in Deutschland legalisiert werden.
C	Eizellspende sollte in Deutschland verboten bleiben.
D	Eizellspenderinnen möchten anonym bleiben.
E	So viele Paare wie möglich sollten sich ihren Kinderwunsch erfüllen können.
F	Kinder sollten die Möglichkeit haben, ihre genetischen Eltern kennenzulernen.
G	In Spanien ist eine anonyme Eizellspende möglich.
H	Eizellspenderinnen sollten eine Aufwandsentschädigung bekommen.



Aufgabe 3: Liegt eine moralische Dimension vor?

Siehst du in folgenden Fällen eine moralische Dimension??

Fall	Siehst du eine moralische Dimension?	
	Ja	Nein
1. Welches Haustier ist das richtige für mich?		
2. Wie sind Bienen organisiert?		
3. Ist es richtig, Organe von Hirntoten zu entnehmen?		
4. Wie verläuft eine HIV-Infektion?		
5. Wie kann ein Bach renaturiert werden?		
6. Welche Zellorganellen kann man im Lichtmikroskop sehen?		
7. Welche Effekte hat Rauchen auf die Lunge?		
8. Welche Tiere sind die natürliche Feinde von Eichhörnchen?		
9. Wie unterscheiden sich die Nervensysteme von Insekten und Säugetieren?		
10. Wie werden embryonale Stammzellen gewonnen?		

Aufgabe 4: Äpfel kaufen

A) Kunde A aus Kassel bevorzugt Äpfel, die biologisch angebaut werden und er möchte lokale Obstbauern unterstützen. Für Kunde B zählt hingegen hauptsächlich der Preis; allerdings mag er keine süßen Äpfel. Welchen der vier Apfelsorten sollten sie jeweils kaufen?

Kunde A: _____

Kunde B: _____

	Apfel 1	Apfel 2	Apfel 3	Apfel 4
Geschmack	süß/fruchtig	sauer	fruchtig/sauer	süß/fruchtig
Anbauart:	konventionell	biologisch	konventionell	biologisch
Anbauort:	Süditalien	Bayern	Neuseeland	Nordhessen
Preis/kg	1,49€	1,99€	1,69€	1,69€

B) Stell dir vor, du wärst Kunde A. Warum kaufst du Äpfel aus biologischem Anbau von lokalen Obstbauern?

Aufgabe 5: Cystische Fibrose

Cystische Fibrose (CF) ist eine autosomal rezessive genetische Krankheit. CF verursacht eine Fehlfunktion der externen Sekretdrüsen, was unter anderem in der Produktion großer Mengen zähen Schleims im Respirationstrakt resultiert. Der Schleim verursacht wiederkehrende Lungenentzündungen, die zu einem Langzeitschaden der Lungen führen. Daher ist die Erkrankung tödlich: Patienten erreichen nur selten das 40. Lebensjahr. Das Gen für CF wurde lokalisiert und Wissenschaftler arbeiten an einer Therapie. Obwohl einige Maßnahmen gefunden wurden, die den Verlauf der Krankheit verlangsamen, sind die Wissenschaftler überzeugt, dass nur mit embryonalen Stammzellen eine dauerhafte Heilung erreicht werden kann. Mehrere Anträge für die Verwendung von embryonalen Stammzellen wurde bereits gestellt, aber die Regierung zögert mit ihrer Entscheidung. Unterdessen leiden die Patienten weiter.

Um an embryonale Stammzellen zu gelangen, müssen menschliche Eizellen mit einem Spermium befruchtet werden. In den frühen Stadien der Ontogenese werden Stammzellen aus der inneren Zellmasse der Blastozyste entnommen. Die Entnahme der Stammzellen resultiert allerdings in der Zerstörung des Embryos, der diese Prozedur nicht überlebt.

A) Welche Werte müssen bei der Nutzung embryonaler Stammzellen bedacht werden?

B) Nenne mögliche Konsequenzen der Legalisierung der Verwendung von embryonalen Stammzellen für pharmazeutische Forschung.

Tom und Rebecca sind beide Träger von CF, leiden aber nicht unter der Krankheit. Sie möchten ein Kind, fürchten aber, dass es an CF leiden könnte. Tom schlägt eine Präimplantationsdiagnostik (PID) vor, aber Rebecca ist sich unsicher. Die PID nutzt die in-vitro Befruchtung. Mehrere Eizellen werden in einer Petrischale befruchtet. Die DNA der Embryos wird auf vererbare Krankheiten untersucht. Nur ein gesunder Embryo wird in die Gebärmutter der Frau implantiert. Die anderen Embryos werden verworfen. Durch PID ist ein Kind ohne genetische Krankheiten sehr wahrscheinlich.

C) Nenne Argumente für PID

Jim und Mary sind verheiratet und Mary ist schwanger. Jim hat einen Bruder, der an CF litt und vor einem Jahr verstorben ist. Mary kennt die Krankheit nicht. Sie haben einen CF Test und beide sind heterozygote Träger der Krankheit. Jetzt denken sie über eine Abtreibung nach.

D) Worin besteht das moralische Problem in diesem Fall?

E) Jim befürwortet eine Abtreibung. Was könnten seine Gründe sein?

F) Sollten sie eine Abtreibung vornehmen? Begründe deine Entscheidung.

Aufgabe 6: Genetisch veränderte Kartoffeln

Eine deutsche Firma hat genetisch veränderte Kartoffeln entwickelt. Die neue Kartoffel ist widerstandsfähiger, wächst schneller und wird von Insekten nicht angegriffen. Um diese Kartoffel anzupflanzen müssen Landwirte das Saatgut und einen speziellen Dünger von der Firma kaufen. Besonders der Dünger ist deutlich teurer als herkömmliche Produkte, aber die Firma behauptet, dass dadurch der Ertrag gesteigert werden könne und sich die Investition schnell auszahle. Verbraucherverbände und Umweltschützer kritisieren die Firma wegen der genetischen Veränderung und betonen, dass diese neue Kartoffel eine Bedrohung für die Natur und die Verbraucher darstellen könnte.

A) Welche der folgenden Werte haben einen Einfluss auf die Situation?

- | | | |
|--|---|--|
| <input type="radio"/> Erhalt der Natur / Naturschutz | <input type="radio"/> Wohlstand | <input type="radio"/> Bildung |
| <input type="radio"/> Verantwortung | <input type="radio"/> Menschenwürde | <input type="radio"/> gesicherte Ernährung |
| <input type="radio"/> Wissenschaftliche Entwicklung | <input type="radio"/> Respekt | <input type="radio"/> Religion/Glaube |
| <input type="radio"/> Gesundheit | <input type="radio"/> Entscheidungsfreiheit | <input type="radio"/> Freiheit |
| <input type="radio"/> Glück | <input type="radio"/> Gerechtigkeit | |

B) Wähle zwei Werte aus, die du für relevant hältst und begründe deine Entscheidung!

TESTHEFT ZUR UNTERSUCHUNG EINER KORRELATION VON KOGNITIVER LEISTUNGSFÄHIGKEIT UND DER LÖSUNG VON BILINGUALEN BIOLOGIEAUFGABEN

Liebe Schülerin, lieber Schüler,

ich bitte dich, diesen Testbogen für meine Examensarbeit auszufüllen.

- Lies dir die Aufgaben genau durch
- Bitte versuche, die Aufgaben so gut wie möglich zu lösen. Wenn du bei einer Aufgabe oder einem Aufgabenteil nicht sicher bist, formuliere bitte trotzdem eine Antwort. In vielen Fällen gibt es keine richtigen oder falschen Antworten, sondern es kommt auf die Begründung an.
- Bitte achte darauf, dass du keine Aufgabe auslässt.
- Wenn dir etwas unklar ist, kannst du mich jederzeit ansprechen.

Deine Antworten werden von mir zwar ausgewertet, aber nicht bewertet. Das bedeutet dass dein Abschneiden bei den Aufgaben keinen Einfluss auf deine Biologienote haben wird. Deine Antworten werden natürlich auch von mir anonym behandelt.

Herzlichen Dank für deine Mitarbeit und viel Spaß!

Zuerst brauche ich einige Angaben von dir:

Alter: _____ Jahre

Geschlecht: männlich weiblich

Jahrgang: 10 11 12 13

Biologieunterricht: Biologie auf Deutsch Biologie bilingual

Seit wann belegst du bilingualen Biologieunterricht? _____

Belegst du auch noch andere bilinguale Fächer? nein ja: _____

Noten/Punkte im letzten Zeugnis: Biologie: _____ Englisch: _____
Deutsch: _____ Mathe: _____

TEIL 1: ERFASSUNG DER KOGNITIVEN LEISTUNGSFÄHIGKEIT.

Dir wird mehrmals hintereinander eine kurze, willkürliche Abfolge von Zahlen gezeigt, **die du dir merken sollst, jedoch ohne sie aufzuschreiben**. Im Anschluss an jede Zahlenfolge werden dir einfache mathematische Rechnungen präsentiert, deren Lösungen du bitte unten notierst. Danach schreibst du bitte die Zahlenfolge auf.

Dieser Vorgang wird mehrmals wiederholt.

Bei diesem Test ist es wichtig, dass du nicht schummelst! Schreibe die Zahlenfolge bitte wirklich erst nach der Bearbeitung der Matheaufgaben auf.

Durchgang 1:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

Durchgang 2:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

Durchgang 3:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

Durchgang 4:

Ergebnisse der Matheaufgaben: _____

Zahlenfolge: _____

TEIL 2: LEISTUNGSTEST BIOLOGIEAUFGABEN.

Task 1: Linda's and Bob's decision

Linda is 16 years old and unintentionally in the 9th week of pregnancy. She and her boyfriend Bob, a guest student from Oklahoma, USA are not quite sure if they should have the baby or not. They don't know enough about the consequences for each decision. Linda's parents are – for religious reasons – strictly against abortion. They want their daughter Linda to give birth to the child and to continue schooling. Bob wants to support Linda but he is afraid of carrying responsibility for the young family on his own. He has no income and hopes to finish school in two years. Bob's parents have only low incomes. Linda and Bob ask their friends to support them in their decision making

A) Define the described dilemma

B) Name actions that are possible solutions for the dilemma

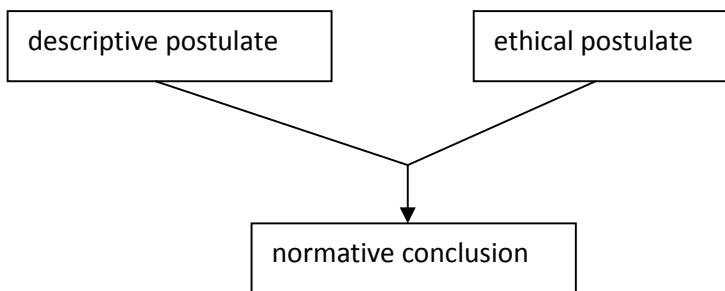
C) Name ethical values that are implied by the respective options.

D) Which decision in this dilemma would you make? Give reasons for it.

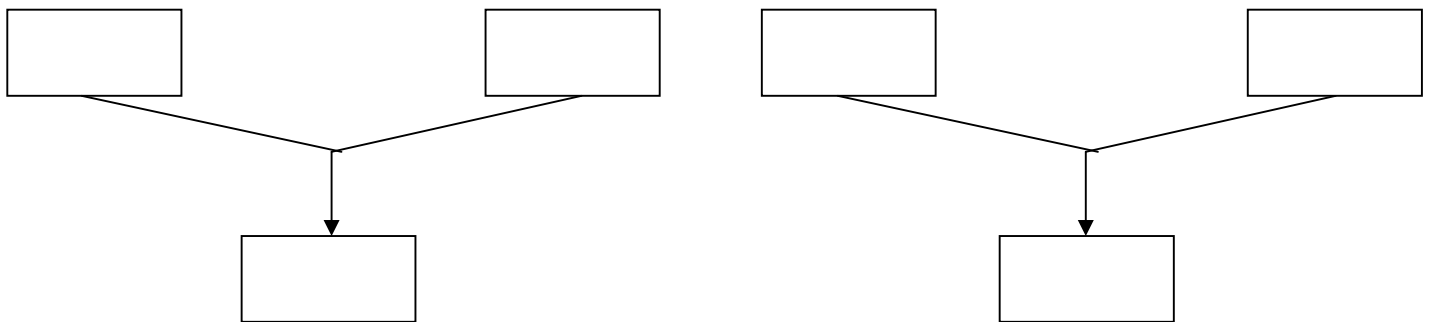
E) List the consequences of your decision in the given case.

Task 2: Egg donation

A) Construct two logical arguments from the statements below. Follow this scheme:



A	Egg donation enables some couples to fulfill their desire for children.
B	Egg donation should be legalized in Germany.
C	Egg donations should remain prohibited in Germany.
D	Egg donors prefer to remain anonymous.
E	As many couples as possible should be able to fulfill their desire for children.
F	Children should have the opportunity to determine their genetic origin.
G	Anonymous egg donation is possible in Spain.
H	Egg donors should be paid an allowance.



Task 3: Ethical dimension or not?

Do you see an ethical dimension in the following cases?

Case	Do you see an ethical dimension	
	Yes	No
1. Which pet is the right one for me?		
2. How are bees organized?		
3. Are we allowed to transplant organs from braindead persons?		
4. What is the course of an HIV infection?		
5. How can a small creek be recultivated?		
6. Which cell organelles are visible in the microscope?		
7. Which effects can smoking have on the lungs?		
8. Which animals are predatory enemies of squirrels?		
9. In what ways do the nervous systems of insects and mammals differ?		
10. How are embryonic stem cells extracted?		

Task 4: Buying Apples [B6]

A) Customer A from Kassel values apple that grew organically and he wants to support local farmers, while the most important aspect for customer B is a low price of the apples. He does not like sweet apples, though. Which of the four types of apples below do they have to purchase?

Customer A: _____

Customer B: _____

	Apple 1	Apple 2	Apple 3	Apple 4
Taste	sweet/fruity	Sour	fruity/sour	sweet/fruity
Growing type:	conventional	Organic	conventional	organic
Grown in:	South Italy	Bavaria	New-Zealand	Northern Hesse
Price/kg	1,49€	1,99€	1,69€	1,69€

B) Imagine you are Customer A. Why do you like buying organic apples from local farmers?

Task 5: Cystic Fibrosis

Cystic Fibrosis (CF) is an autosomal recessive genetic trait. CF causes a deficient functioning of the external secretion glands that is pronounced (among other things) in the production of large quantities of mucus in the respiratory tracts. The mucus causes recurrent lung infections which add to the long-term damage to the lungs. The disease is therefore lethal: patients rarely survive past the age of 40. The gene responsible for CF has been located and scientists are now working on a treatment. Although some treatments have been found that slow the course of the disease, scientists believe that only the usage of embryonic stem cells could provide a permanent cure. Several petitions to use embryonic stem cells were already brought under way, but the government hesitates to take a decision. Meanwhile, patients keep suffering.

To gain embryonic stem cells, human eggs have to be inseminated with a sperm. In the early stages of the ontogenesis, embryonic stem cells are derived from the inner cell mass of the blastocyst. Isolating these cells, however, results in the destruction of the embryo, who does not survive this procedure.

A) Which values have to be taken into consideration to decide about the usage of embryonic stem cells?

B) Name possible consequences of the legalization of the use of embryonic stem cells for pharmaceutical matters.

Tom and Rebecca are both carriers of CF but do not suffer from the disease, which is autosomal recessive genetic trait. They want to have a baby but are afraid that it might suffer from CF. Tom suggest Pre-Implantation Genetic Diagnosis (PGD) but Rebecca is unsure about it. The pre-implantation diagnostic makes use of an in-vitro fertilization. Several eggs are impregnated in a petri dish. The developing embryos' DNA is analyzed for diseases. Only a healthy embryo is implanted into the woman's uterus, the other fertilized eggs are discarded. Thus, this method makes it highly likely that the baby will be free of the disease under consideration

C) Name arguments in favor of PGD _____

Jim and Mary are married and Mary got pregnant. Jim has a brother who suffered from CF and died a year ago. Mary does not know the disease. They had a CF test and both are homozygous for CF. Now they contemplate whether or not they should have an abortion.

D) What is the moral problem under consideration? _____

E) Jim leans towards having an abortion. What could be his reasons?

F) Should they have the abortion? Offer arguments for your position.

Task 6: Genetically modified potatoes.

A German company designed a genetically modified potato. The new potato is more durable, grows faster and is unharmed by insects and bugs. To grow these potatoes, farmers need to buy seeds and special fertilizer from the company. The fertilizer in particular is more expensive than common products, but the company claims that the harvest is bigger and that the investment pays off. Customer interest groups and environmentalists criticize the company for modifying products genetically and point out that the potatoes could be dangerous for the wildlife and customers.

A) Which of the following values have an impact on this situation?

- | | | |
|--|---|--|
| <input type="radio"/> Environmental Preservation | <input type="radio"/> Prosperity | <input type="radio"/> Education |
| <input type="radio"/> Responsibility | <input type="radio"/> Human Dignity | <input type="radio"/> Nutritional Security |
| <input type="radio"/> Advancement of Science | <input type="radio"/> Respect | <input type="radio"/> Religion/Beliefs |
| <input type="radio"/> Health | <input type="radio"/> Freedom of Choice | <input type="radio"/> Freedom |
| <input type="radio"/> Luck/Happiness | <input type="radio"/> Justice | |

B) Pick two of the values you consider relevant reason your decision:

TEIL 3: FRAGEBOGEN ZUM BILINGUALEN BIOLOGIEUNTERRICHT

Dieser Fragebogen behandelt dein Empfinden des Biologieunterrichts allgemein. Bitte antworte so ehrlich wie möglich und mach pro Frage immer nur ein Kreuz!

		Stimmt genau	Stimmt fast	Weder noch	Stimmt kaum	Stimmt gar nicht
1	In Biologie bin ich sehr motiviert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	Biologie finde ich besser als andere Fächer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	In Biologie fühle ich mich aufgrund der englischen Sprache oftmals überfordert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	In Biologie fühle ich mich aufgrund des Inhalts oftmals überfordert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	Ich würde Biologie wieder als bilinguales Fach wählen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	Ich bereue meine Entscheidung, Biologie bilingual gewählt zu haben	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	Ich würde im Biologieunterricht gerne mehr auf Deutsch sprechen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	Ich würde im Biologieunterricht gerne ganz auf Deutsch verzichten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Vielen Dank für deine Hilfe!

Anhang 3: Erwartungshorizont und Codierungsanleitung

Fette Markierung: Erwartungen/mögliche Antworten und Codierungsanleitung

Kursive Markierung: Teilkompetenz, die von einer Aufgabe hauptsächlich abgefragt wird

TEIL 1: ERFASSUNG DER KOGNITIVEN LEISTUNGSFÄHIGKEIT.

Dir wird mehrmals hintereinander eine kurze, willkürliche Abfolge von Zahlen gezeigt, **die du dir merken sollst, jedoch ohne sie aufzuschreiben**. Im Anschluss an jede Zahlenfolge werden dir einfache mathematische Rechnungen präsentiert, deren Lösungen du bitte unten notierst. Danach schreibst du bitte die Zahlenfolge auf.

Dieser Vorgang wird mehrmals wiederholt.

Bei diesem Test ist es wichtig, dass du nicht schummelst! Schreibe die Zahlenfolge bitte wirklich erst nach der Bearbeitung der Matheaufgaben auf.

Durchgang 1:

Zahlenfolge: 630932

Durchgang 2:

Zahlenfolge: 3768148

Durchgang 3:

Zahlenfolge: 99807325

Durchgang 4:

Zahlenfolge: 646351507

FULL CREDIT (1) = Ziffernreihenfolge richtig

NO CREDIT = Ziffernfolge falsch, Anzahl der Fehler irrelevant

TEIL 2: LEISTUNGSTEST BIOLOGIEAUFGABEN.

Task 1: Linda's and Bob's decision

[...]

A) Define the described dilemma

Wahrnehmen und Bewusstmachen ethischer Relevanz

Linda und Bob müssen sich entscheiden, ob sie das Baby bekommen möchten, oder ob sie ihre eigenen Lebensträume verwirklichen. Beides ist nicht gleichzeitig möglich, eine Option muss hintenan gestellt werden. Das Leben des ungeborenen Babys steht im Konflikt mit der Verwirklichung der Pläne und Träume des jugendlichen Paares. Außerdem sind verschiedene weitere ethische Werte betroffen, die mitunter nicht miteinander vereinbar sind, z.B. die Religiosität der Eltern

Niveau 0: keine, oder falsche Beschreibung

Niveau I: Deskriptiv-pragmatische Problemwahrnehmung

Niveau II: Deskriptive Problemwahrnehmung mit Bewusstmachung des ethisch-moralischem Problems, möglicherweise emotionale Färbung

Niveau III: Objektive Erkenntnis des moralisch-ethischen Wertkonflikts

B) Name actions that are possible solutions for the dilemma

Beurteilen

Optionen sind z.B.: Abtreibung, Freigabe zur Adoption, Kind von Eltern/Großeltern aufziehen lassen, Kind bekommen und Familie gründen.

Niveau 0: keine Option wird genannt

Niveau I: 2 Optionen

Niveau II: mehr als 2 Optionen

C) Name ethical values that are implied by the respective options.

Ethisches Basiswissen

Werte: Leben, Recht auf Glück, Recht auf Selbstbestimmung, Wahlfreiheit, Religion, Verantwortung, ...

Niveau 0: Der Schüler nennt keinen relevanten Wert

Niveau I: Der Schüler nennt nur einen relevanten Wert
Niveau II: Der Schüler nennt zwei relevante Werte
Niveau III: Der Schüler nennt mehr als zwei relevante Werte

D) Which decision in this dilemma would you make? Give reasons for it.

Urteilen / Argumentieren

Entscheidung für einer der Optionen aus Aufgabenteil B. Die Anzahl, Bedeutung und Sinnhaftigkeit der Argumente, die die Entscheidung bekräftigen macht den Unterschied zwischen Niveau II und III. Unklare Aussagen oder ein Abwägen von Entscheidungsmöglichkeiten wird mit Niveau I bewertet.

Niveau 0 = es findet keine Aussage statt.

Niveau I = Der Schüler macht eine zusammenhangslose Aussage, die zudem keine Begründung enthält.

Niveau II = Der Schüler bezieht klar Stellung, begründet seine Meinung aber nicht oder kaum

Niveau III = Der Schüler bezieht klar Stellung und begründet seine Meinung mit einem logisch konsistenten Argument

E) List the consequences of your decision in the given case.

Folgenreflexion

**Primärfolgen: Geburt/Tod (Baby), Mutterschaft (Linda), Schulabbruch (Bob)
Sekundärfolgen: Familiengründung, Schuldgefühle, zerrüttetes Verhältnis zu Lindas Eltern, Ermordung eines Menschen,...**

Niveau 0: keine Folgen werden gelistet

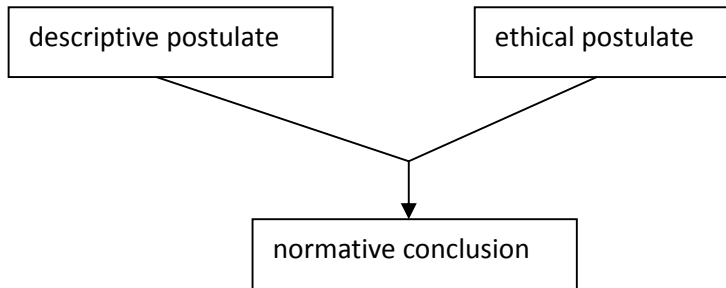
Niveau I: Primärfolgen und Sekundärfolgen und deren Ausdifferenzierung für unmittelbar Betroffene

Niveau II: Primärfolgen und Sekundärfolgen und deren Ausdifferenzierungen für unmittelbar und mittelbar Betroffene, Langzeitfolgen

Task 2: Egg donation

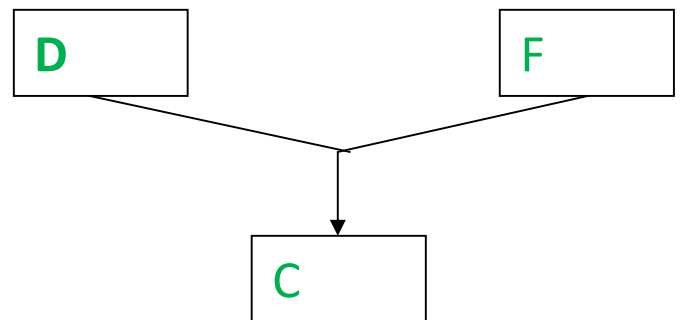
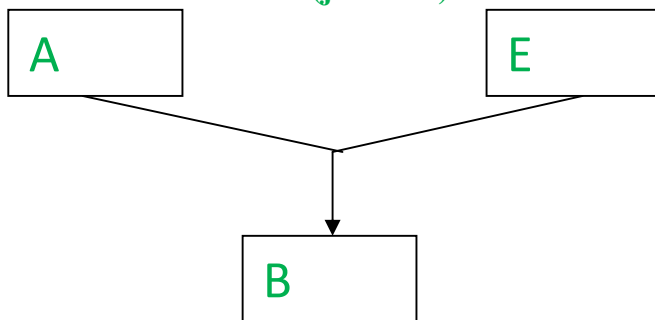
Ethisches Basiswissen

A) Construct two logical arguments from the statements below. Follow this scheme:



A	Egg donation enables some couples to fulfill their desire for children.
B	Egg donation should be legalized in Germany.
C	Egg donations should remain prohibited in Germany.
D	Egg donors prefer to remain anonymous.
E	As many couples as possible should be able to fulfill their desire for children.
F	Children should have the opportunity to determine their genetic origin.
G	Anonymous egg donation is possible in Spain.
H	Egg donors should be paid an allowance.

FULL CREDIT (jeweils 1) =



NO CREDIT = Alle anderen Kombinationen

Task 3: Ethical dimension or not?

Wahrnehmen und Bewusstmachen ethischer Relevanz

Do you see an ethical dimension in the following cases?

Case	Do you see an ethical dimension	
	Yes	No
1. Which pet is the right one for me?	X	
2. How are bees organized?		X
3. Are we allowed to transplant organs from braindead persons?	X	
4. What is the course of an HIV infection?		X
5. How can a small creek be recultivated?		X
6. Which cell organelles are visible in the microscope?		X
7. Which effects can smoking have on the lungs?		X
8. Which animals are predatory enemies of squirrels?		X
9. In what ways do the nervous systems of insects and mammals differ?		X
10. How are embryonic stem cells extracted?	X	

FULL CREDIT = alle richtig

HALF CREDIT = 1-3 falsch

NO CREDIT = >3 falsch

Task 4: Buying Apples

Beurteilen

A) Customer A from Kassel values apple that grew organically and he wants to support local farmers, while the most important aspect for customer B is a low price of the apples. He does not like sweet apples, though. Which of the four types of apples below do they have to purchase?

Customer A: Apple 4

Customer B: Apple 3

FULL CREDIT = Apple 4 (A) Apple 3 (B)

NO CREDIT = andere Apfelsorten, mehrere Apfelsorten pro Kunde

	Apple 1	Apple 2	Apple 3	Apple 4
Taste	sweet/fruity	Sour	fruity/sour	sweet/fruity
Growing type:	Conventional	Organic	conventional	Organic
Grown in:	South Italy	Bavaria	New-Zealand	Northern Hesse
Price/kg	1,49€	1,99€	1,69€	1,69€

B) Imagine you are Customer A. Why do you like buying organic apples from local farmers?

Perspektivwechsel

Argumente: Unterstützung lokaler Händler/Landwirte, Gesundheit, Einsparen von CO₂ durch kurze Transportwege, Verzicht auf Pestizide und Chemikalien beim Anbau der Äpfel, ...

Niveau 0 = Der Schüler nennt keine Argumente.

Niveau I = Der Schüler nennt ein Argument.

Niveau II = Der Schüler nennt zwei Argumente

Niveau III = Der Schüler nennt drei oder mehr Argumente

Task 5: Cystic Fibrosis

Cystic Fibrosis (CF) is an autosomal recessive genetic trait. CF causes a deficient functioning of the external secretion glands that is pronounced (among other things) in the production of large quantities of mucus in the respiratory tracts. The mucus causes recurrent lung infections which add to the long-term damage to the lungs. The disease is therefore lethal: patients rarely survive past the age of 40. The gene responsible for CF has been located and scientists are now working on a treatment. Although some treatments have been found that slow the course of the disease, scientists believe that only the usage of embryonic stem cells could provide a permanent cure. Several petitions to use embryonic stem cells were already brought under way, but the government hesitates to take a decision. Meanwhile, patients keep suffering.

To gain embryonic stem cells, human eggs have to be inseminated with a sperm. In the early stages of the ontogenesis, embryonic stem cells are derived from the inner cell mass of the blastocyst. Isolating these cells, however, results in the destruction of the embryo, who does not survive this procedure.

A) Which values have to be taken into consideration to decide about the usage of embryonic stem cells?

Ethisches Basiswissen

Werte: Fortschritt der Wissenschaft, Menschenwürde, Religion, Gesundheit, Menschenleben,...

Niveau 0: kein relevanter Wert wird genannt

Niveau I: ein relevanter Wert wird genannt

Niveau II: zwei relevante Werte werden genannt

Niveau III: mehr als zwei relevante Werte werden genannt

B) Name possible consequences of the legalization of the use of embryonic stem cells for pharmaceutical matters.

Beurteilen

Folgen: Heilbarkeit diverser Krankheiten, medizinischer Fortschritt, Leben können gerettet werden, Leiden wird gelindert, Embryos werden abgetötet, Wissenschaftler werden gottgleiche Figuren, Menschenopfer für die Wissenschaft, ethische Bedenken der Spender

Niveau 0: keine Folgen werden gelistet

Niveau I: Primärfolgen und deren Ausdifferenzierung für unmittelbar Betroffene

Niveau II: Primärfolgen und Sekundärfolgen für unmittelbar und mittelbar Betroffene, Langzeitfolgen

Niveau III: Gesellschaftliche und persönliche Folgen werden genannt

Tom and Rebecca are both carriers of CF but do not suffer from the disease, which is autosomal recessive genetic trait. They want to have a baby but are afraid that it might suffer from CF. Tom suggest Pre-Implantation Genetic Diagnosis (PGD) but Rebecca is unsure about it. The pre-implantation diagnostic makes use of an in-vitro fertilization. Several eggs are impregnated in a petri dish. The developing embryos' DNA is analyzed for diseases. Only a healthy embryo is implanted into the woman's uterus, the other fertilized eggs are discarded. Thus, this method makes it highly likely that the baby will be free of the disease under consideration

C) Name arguments in favor of PGD

Argumentieren

Argumente: nur gesunde Kinder können geboren werden, Sieg gegen Erbkrankheiten, Glück der Eltern und Kinder kann gesichert werden, ...

Niveau 0 = der Schüler nennt keine Argumente, oder Argumente gegen PID

Niveau I = Der Schüler nennt ein Argument für PID

Niveau II = Der Schüler nennt zwei oder mehr Argumente für PID

Jim and Mary are married and Mary got pregnant. Jim has a brother who suffered from CF and died a year ago. Mary does not know the disease. They had a CF test and both are homozygous for CF. Now they contemplate whether or not they should have an abortion.

D) What is the moral problem under consideration?

Wahrnehmen und Bewusstmachen moralischer Relevanz

Dilemma: Die Chance, dass das Baby an CF erkrankt ist sehr hoch. Daher besteht auch ein hohes Risiko für ein Leben voller Leiden für Kind und Eltern. Andererseits könnte auch ein vollkommen gesundes Baby abgetrieben werden und das Paar müsste sich vom Kinderwunsch verabschieden.

Niveau 0: keine, oder falsche Beschreibung

Niveau I: Deskriptiv-pragmatische Problemwahrnehmung

Niveau II: Deskriptive Problemwahrnehmung mit Bewusstmachung des ethisch-moralischem Problems, möglicherweise emotionale Färbung

Niveau III: Objektive Erkenntnis des moralisch-ethischen Wertkonflikts

E) Jim leans towards having an abortion. What could be his reasons?

Perspektivwechsel

Gründe: Erfahrungen mit Krankheit und Tod des Bruders, Angst vor Verantwortung, Angst vor Leiden und Tod des Kindes, ...

Niveau 0 = keine Gründe werden genannt.

Niveau I = ein Grund wird genannt

Niveau II = zwei Gründe werden genannt

Niveau III = mehr als zwei Gründe werden genannt

F) Should they have the abortion? Offer arguments for your position.

Urteilen / Bewusstmachen der eigenen Einstellung / Argumentieren

Klare Begründung für oder gegen eine Abtreibung und Nennung von Argumenten für die jeweilige Entscheidung

Niveau 0 = der Schüler macht keine Aussage oder trifft keine Entscheidung

Niveau I = Der Schüler macht eine zusammenhangslose Aussage, die zudem keine Begründung enthält.

Niveau II = Der Schüler bezieht klar Stellung, begründet seine Meinung aber nicht oder kaum

Niveau III = Der Schüler bezieht klar Stellung und begründet seine Meinung mit einem logisch konsistenten Argument

Task 6: Genetically modified potatoes.

AUFGABE WIRD NICHT AUSGEWERTET