

Pädagogische Hochschule Freiburg

Hochschulschrift

Entwicklung und Evaluation von Methoden zum Umgang mit Heterogenität im Biologieunterricht

Zur Erlangung des akademischen Grades Dr. phil.

Institut für Biologie und ihre Didaktik

Wolfgang Feller

Dekanin: Prof. Dr. Ulrike Spörhase

Gutachter:

1. Prof. Dr. Ulrike Spörhase
2. Prof. Dr. Silke Mikelskis-Seifert
3. Prof. Dr. Marco Oetken

Datum der Prüfung: 13.10.2015

SUMMARY

Introduction: To cope with heterogeneity of performance, teachers have to differentiate within their class, however internal differentiation hardly takes place in biology classes so far.

Objectives: The main objectives of this study are to find out (a) which approaches of internal differentiation to teaching biology are suitable from the didactical point of view? Which of these approaches do biology teachers accept based on their experience and knowledge? (b) Does this approach, a differentiating representation of nearly identical biological content (in this case expert-concept-maps versus plain text), have an effect on the performance of underachieving students? (c) Do any gender differences occur in connection to the intervention? (d) Does study time influence the results of students? (e) How are the learning processes of students with the learning material and where do difficulties in understanding arise?

Methods: (a) As an initial step, practical approaches to internal differentiation have been discussed and actively tried out with twelve teachers in a symbiotic setting (all women with different levels of job experience) during a teacher training with ten meetings over a period of 1.5 years. The data of this step were collected through participant observation. The second step included the development of an approach to internal differentiation. (b) The effect of this approach (expert-concept-map versus text) on underachieving students was tested through dyadic interviewing with 112 students from schools (Realschule) in the southern Black Forrest region of the German federal state of Baden-Württemberg (62 ♀ / 50 ♂, age 12 to 14). The data for the developed internal differentiation approach were collected with a paper-questionnaire and analyzed by conducting a two-way ANOVA. (c) Gender related differences were also analyzed through quantitative methods (two-way ANOVA, linear regression and t-tests) and additionally confirmed by qualitative data through video graphic methods of the interviews. The same combinations of qualitative and quantitative data were conducted for the objectives (d) and (e).

Results: (a) The participant observation of our teachers showed five points according to possible ways of internal differentiation: First, the method should be assigned by the teacher to certain students (not picked by students themselves) and should support underachieving students explicitly. Second, the method should not widen the knowledge gap between students but keep them on the same level so they can write the same tests. Third, the method should be plain and simple, with task instructions that are easy to understand, preferably with an integrated task solution or worked-out example that students have the chance to correct their answers themselves. Forth, the

method should not require much additional time and material for the preparation and follow-up. Fifth, the method material and instruction should make use of as little written text as possible. (b) The examination of the effectiveness of our approach showed that expert-concept-maps are good for the transmission of content for all students. (c) There were no measurable differences between genders through the intervention, but other gender differences occurred. The disparities between sexes first showed up when the submitted school grades of students were related to their actual performance in this study, both qualitatively and quantitatively. (e) Female students didn't perform as well as they were expected to by their school grade. Furthermore they underperformed especially in tasks of application of knowledge and transfer. (d) Regarding the performance of students in relation to study time, it turned out that learning time of underachievers made a significant difference in comparison to achievers.

Conclusion: The results of this study not only indicated that expert-concept-maps could be used for the transmission of content for all students, they also proofed that learning time is a powerful tool for differentiation. In addition, the results provide approaches to further developments of internal differentiation in biology classes and give indications on the necessity of a more gender-sensitive teaching.

ZUSAMMENFASSUNG

Hintergrund: Um der Leistungsheterogenität in ihren Klassen gerecht zu werden, müssen Lehrkräfte (binnen-)differenzieren. Dies findet im Biologieunterricht bisher jedoch kaum statt.

Ziele: Vor diesem Hintergrund verfolgt diese Studie folgende Ziele: (a) Welche Ansätze innerer Differenzierung eignen sich aus fachdidaktischer Sicht für die Vermittlung biologischer Inhalte? Welche dieser Ansätze nehmen Lehrkräfte aufgrund ihres Erfahrungswissens an? (b) Hat eine differenzierende Darstellungsform annähernd identischer Lehrinhalte über Experten-Concept-Maps oder Text Auswirkungen auf das Wissen und Können bei Lernenden mit unterschiedlichen Leistungsniveaus? (c) Treten bei den eingesetzten Differenzierungsmaßnahmen geschlechtsspezifische Unterschiede auf? (d) Welchen Einfluss hat die Lernzeit auf das Ergebnis der Schüler? (e) Wie gestaltet sich der Lernprozess der Schüler mit den differenzierten Darstellungsformen und der Lernaufgaben und wo treten dabei Schwierigkeiten auf?

Methodik: (a) In einem ersten Schritt wurden praktische Ansätze zu Binnendifferenzierung während einer Lehrerfortbildung in einem symbiotischen Setting mit zwölf Lehrkräften diskutiert und ausprobiert. Die Fortbildung erstreckte sich über einen Zeitraum von 1,5 Jahren mit insgesamt zehn Sitzungen. Die Teilnehmer setzten sich aus Lehrerinnen zwischen 26 und 57 Jahren mit unterschiedlicher Berufserfahrung zusammen. Die Daten aus der Fortbildung wurden mittels der Methode der teilnehmenden Beobachtung erhoben. (b) Die Wirkung des kooperativ entwickelten Ansatzes wurde in Dyaden-Interviews mit 112 Realschülern aus dem südlichen Schwarzwald in Baden-Württemberg überprüft (62 ♀ / ♂ 50, Alter 12 bis 14). Hierfür wurden Daten mit einem Fragebogen erhoben und mittels einer zweifaktoriellen Varianzanalyse ohne Messwiederholung untersucht (two-way ANOVA). (c) Darüber hinaus wurden die Daten durch lineare Regressionen und t-Tests auf geschlechtsspezifische Unterschiede hin analysiert und durch qualitative Daten aus der Videoanalyse der Dyaden-Interviews ergänzt. Auf die gleiche Weise wurden die Ziele (d) und (e) überprüft.

Ergebnisse: (a) Die teilnehmende Beobachtung erbrachte fünf Punkte für das Gelingen eines Ansatzes zur Binnendifferenzierung im Biologieunterricht: Erstens sollte der Differenzierungsansatz von Lehrkräften den jeweiligen Schülern zugewiesen und nicht von diesen selbst ausgewählt werden. Ferner sollte er explizit leistungsschwächere Schüler unterstützen. Zweitens sollte das Verfahren die Wissenslücken zwischen den Schülern nicht erweitern sondern die Lernenden möglichst auf einem Level halten, damit diese zur Leistungsermittlung die gleichen Tests schreiben können. Drittens sollte das Verfahren schlicht und einfach in der Durchführung sein, mit leicht zu ver-

stehenden Aufgabenstellungen und vorzugsweise mit einer integrierten Musterlösung oder Beispiel-Antwort, damit die Schüler sich selbst korrigieren können. Viertens sollte die Methode nicht zu viel zusätzliche Zeit und Materialien in der Vor- und Nachbereitung beanspruchen und fünftens von so wenig geschriebenem Text wie möglich Gebrauch machen. (b) Bei der Überprüfung dieses Ansatzes hat sich gezeigt, dass sich Experten-Concept-Maps für die Inhaltsvermittlung bei allen Schülern eignen. (c) Bei der Überprüfung der Intervention traten keine messbaren Unterschiede zwischen den Geschlechtern auf, jedoch ergaben sich unerwartete Ergebnisse bei der Analyse der Daten. Diese traten auf, als die von den Schülern in dieser Studie erbrachten Leistungen in Relation zu ihren eingereichten Schulnoten gesetzt wurden. (e) Hierbei zeigte sich qualitativ wie quantitativ, dass Mädchen nicht so gut abgeschnitten haben wie es aufgrund ihrer Schulnoten zu erwarten gewesen wäre. Dies machte sich besonders bei Aufgaben der Anwendung und des Transfers von Wissen bemerkbar. (d) In Bezug auf Schülerleistung und Zeit kam heraus, dass die Lernzeit bei leistungsschwächeren Schülern im Vergleich zu leistungstärkeren Schülern einen signifikanten Unterschied macht.

Fazit: Die Ergebnisse dieser Studie deuten darauf hin, dass Experten-Concept-Maps für die Inhaltsvermittlung bei allen Schülern eingesetzt werden sollten und die Lernzeit eine effektive und wirksame Differenzierungsmöglichkeit darstellt. Zudem liefern die Ergebnisse Ansätze für weitere Entwicklungen von Differenzierungsansätzen im Biologieunterricht und geben Hinweise auf die Notwendigkeit eines mehr gendersensiblen Unterrichts.

DANKSAGUNG

Zu Beginn möchte ich mich bei all denjenigen bedanken, die mich während der Doktorarbeit unterstützt und motiviert haben.

Ich möchte mich bei Ulrike Spörhase und Silke Mikelskis-Seifert für ihre Betreuung, ihren Einsatz und ihre Zeit bedanken. Dies gilt sowohl für die schönen und ganz besonders für die weniger schönen Phasen meiner Promotion, in denen ich mich stets auf ihren Beistand verlassen konnte.

Ich danke allen Teilnehmenden der Lehrerfortbildung für die gewinnbringende Zusammenarbeit. Gerade in den gemeinsamen Arbeitsphasen und Diskussionen hat mich die Offenheit und Ehrlichkeit beeindruckt. Damit meine ich nicht nur den fachlichen Austausch, sondern auch die gegenseitige Unterstützung und den kollegialen Beistand.

Ferner danke ich dem staatlichen Schulamt Offenburg für dessen Unterstützung und hierbei namentlich Wolfgang Zink. Ich habe mich über die Genehmigung der Studie durch das Kultusministerium von Baden-Württemberg gefreut, die den Schulen die Teilnahme an diesem Projekt ermöglichte. Den jeweiligen Schulleitungen danke ich für die unkomplizierte Aufnahme in den Schulbetrieb und für die Bereitstellung eines extra Raumes, was nicht immer einfach war. Außerdem danke ich für die Gespräche mit den Schulleitungen, die sich bei meiner Vorstellung an den Schulen beziehungsweise vor und nach den Interviewtagen ergeben haben. Die Gespräche eröffneten mir eine Perspektive auf Binnendifferenzierung, welche mir bisher verborgen war.

Ein besonderer Dank geht an die Lehrkräfte, die sich dazu bereit erklärt haben, mich mit ihren Klassen bei dem Projekt zu unterstützen. Das gezeigte Engagement war bemerkenswert und nicht selbstverständlich.

An dieser Stelle auch ein Dankschön an die teilnehmenden Schüler. Für viele Schüler war die Interviewsituation vor laufender Kamera am Anfang ungewohnt. Einigen Schülern hat es aber auch zunehmend Spaß gemacht. Beides kann man in den Videos deutlich sehen.

Genauso möchte ich mich bei meiner Doktorandencrew bedanken. Die Treffen waren nicht nur eine fachliche Bereicherung, sondern auch ein Mittel der Psychohygiene. Ganz besonders hat mir die Unterstützung geholfen, wenn ich den Wald vor Bäumen nicht mehr gesehen habe oder ich vom Tausendstel ins Millionstel abgedriftet bin.

Ein weniger fachliches, aber umso herzlicheres Dankeschön geht an meine Familie, meine Freunde und an Juliane. Ich habe mich während der Promotion zunehmend

seltener gemeldet, war manchmal gereizt, häufig müde und oft neben der Spur. Ihr habt mich ertragen, mich unterstützt und zu mir gehalten. Dafür danke ich euch. Ohne euren Beistand wäre ich nicht so weit gekommen. Ich gelobe Besserung und freue mich auf eine gemeinsame Zukunft mit euch.

INHALTSVERZEICHNIS

Summary	I
Zusammenfassung	III
Danksagung	V
Inhaltsverzeichnis	VII
Abkürzungen	XI
Abbildungsverzeichnis	XII
Tabellenverzeichnis	XVI
1 Einleitung und Zielsetzung der Arbeit	1
2 Problemstellung und Aufbau der Arbeit	3
3 Theoretische Fundierung relevanter Themen und Inhalte	7
3.1 Methodisches Vorgehen.....	8
3.2 Differenzierung im Unterricht	11
3.2.1 Allgemeine Aspekte von Differenzierung in der Schule	14
3.2.2 Voraussetzungen für differenzierenden Unterricht.....	19
3.2.3 Kriterien von Differenzierung im Unterricht.....	23
3.2.4 Veränderungen der Lehrerrolle durch Binnendifferenzierung	25
3.2.5 Zusammenfassung.....	33
3.3 Lehren und Lernen mit Repräsentationen.....	36
3.3.1 Cognitive Load Theory nach Chandler und Sweller.....	37
3.3.2 Theorie zum Lernen mit Medien nach Mayer	39
3.3.3 Gütekriterien für Lehrmaterialien nach Horz und Schnotz	41
3.3.4 Externe Repräsentationen (Medien) im Biologieunterricht.....	43
3.3.5 Zusammenfassung.....	44
3.4 Concept-Maps als Lerngelegenheiten für den Unterricht.....	45
3.4.1 Lernen mit Concept-Maps	50
3.4.2 Experten-Concept-Maps, eine Sonderform	51
3.4.3 Erwiesene Lernwirksamkeit von Experten-Concept-Maps	52
3.4.4 Experten-Concept-Map im Vergleich zu Text.....	52
3.4.5 Experten-Concept-Maps als Differenzierungsmethode.....	53
3.4.6 Zusammenfassung.....	54
3.5 Aufgaben im Unterricht	56
3.5.1 Formale Aspekte von Aufgaben	57
3.5.2 Funktionen von Aufgaben im Unterricht.....	63
3.5.3 Unterscheidung zwischen Test- und Lernaufgaben.....	67
3.5.4 Qualitätskriterien für Lernaufgaben.....	71

3.5.5	Schwierigkeitsgrade von Aufgaben	76
3.5.6	Überblick der Perspektiven auf den Schwierigkeitsgrad	84
3.6	Differenzierung durch Lernaufgaben.....	86
4	Erarbeitung von Materialien in Kooperation mit Lehrkräften	89
4.1	Einleitung	89
4.2	Methodisches Vorgehen bei der Lehrerfortbildung	92
4.3	Ergebnisse der Lehrerfortbildung	97
4.3.1	Warum findet Binnendifferenzierung bisher kaum statt?	97
4.3.2	Welche Differenzierungsmaßnahmen bevorzugen Lehrkräfte?	99
4.3.3	Rückmeldung der Lehrenden auf die Fortbildung	101
4.4	Zusammenfassung.....	103
5	Allgemeine methodische Vorgehensweisen und eingesetzte Materialien.....	105
5.1	Einleitung	105
5.2	Rekrutierung und Beschreibung der Testpersonen	105
5.3	Datenschutz (Fragebögen und Schülerinterviews)	107
5.4	Einteilung der Gruppen.....	108
5.5	Formative Evaluation des Input und der Aufgaben	108
5.6	Vorstellung des differenzierten Input	110
5.7	Vorstellung der Lernaufgaben 1 bis 4.....	112
5.8	Ablauf der Interviews	113
5.9	Methoden zur Auswertung der Daten	116
6	Empirische Untersuchung	125
6.1	Einleitung	125
6.2	Erkenntnisse aus der Erprobung der Materialien.....	126
6.2.1	Schüler verfolgen individuelle Lernwege	126
6.2.2	Das Augenmodell fördert Verstehen und Kommunikation	126
6.3	Experten-Concept-Maps im Vergleich zu Text	127
6.3.1	Alle Schüler profitieren von Experten-Concept-Maps	127
6.3.2	Leistungsschwächere Schüler profitieren marginal	128
6.4	Genderaspekte.....	129
6.4.1	Mädchen haben bessere Noten, aber keine besseren Ergebnisse.....	129
6.4.2	Beide Geschlechter profitieren von Experten-Concept-Maps	129
6.4.3	Das Geschlecht hat Einfluss auf Schulnoten	131
6.5	Ergebnisse zum Faktor Zeit	132
6.5.1	Der Zeitbedarf ist individuell verschieden.....	132
6.5.2	Leistungsschwächere Schüler benötigen mehr Unterstützung	133
6.5.3	Leistungsschwächere Schüler erzielen bessere Ergebnisse, wenn sie sich Zeit nehmen.....	136

6.5.4	Zeitbedarf bei den Lernaufgaben 1 bis 4	139
6.6	Einsatz der Hilfen bei den Lernaufgaben.....	144
6.6.1	Bei Aufgabe 4 wurden die meisten Hilfen benötigt	144
6.6.2	Qualität der benötigten Hilfe	145
6.7	Schwierigkeiten der Schüler beim Lösen der Aufgaben	146
6.8	Weitere Fragestellungen	154
6.8.1	Einschätzung der Lernaufgaben durch die Schüler	154
6.8.2	Was eine Aufgabe nach Schülermeinung einfach macht.....	156
6.8.3	Was eine Aufgabe nach Schülermeinung schwierig macht.....	157
6.8.4	Was eine Aufgabe nach Schülermeinung interessant macht	158
6.8.5	Qualität der Schülerantworten	161
6.8.6	Schwierigkeiten der Schüler bei der Interpretation der Abbildungen	167
6.8.7	Einfluss des Inputs auf die Antwortqualität bei Reproduktionsaufgaben	169
6.9	Zusammenfassung.....	171
7	Diskussion.....	175
7.1	Zu der Lehrerfortbildung	175
7.2	Zu der formativen Evaluation der Materialien	176
7.2.1	Schüler zeigten unterschiedliche Lernwege.....	176
7.2.2	Modelle fördern Verstehen	177
7.3	Zu den Ergebnissen von Experten-Concept-Maps und Text.....	177
7.3.1	Alle Schüler profitieren von Experten-Concept-Maps bei reiner Reproduktion.....	177
7.3.2	Leistungsschwächere Schüler profitieren marginal von Experten- Concept-Maps	178
7.4	Zu den Genderaspekten.....	179
7.5	Zu den Ergebnissen zum Faktor Zeit.....	183
7.5.1	Zeit als wirksamer Differenzierungsfaktor	183
7.5.2	Leistungsschwächere Schüler zu fördern erfordert mehr Zeit von Lehrkräften.....	184
7.6	Zu dem Einsatz der Hilfen bei den Lernaufgaben	184
7.6.1	Gestufte Hilfen bilden Verstehensschwierigkeiten ab und machen Unterschiede zwischen Schülern deutlich	184
7.6.2	Ein Viertel aller Schüler benötigt nicht-fachliche Unterstützung.....	188
7.7	Der Schwierigkeitsgrad von Aufgaben ist planbar	189
7.8	Der Schwierigkeitsgrad einer Aufgabe fördert Interesse.....	189
7.9	Schwierigkeiten der Schüler bei den Abbildungen.....	190
7.10	Limitationen der Studie.....	191
7.11	Abschließende Bemerkungen	192

8 Quellenverzeichnis.....	195
9 Anhang.....	219
9.1 Darstellung der zusätzlich eingesetzten Materialien	219
9.1.1 Input: Fachtext	221
9.1.2 Input: Experten-Concept-Map	222
9.1.3 Multiple-Choice-Test zu dem Input.....	223
9.1.4 Input: Abbildungen zum Auge.....	224
9.1.5 Multiple-Choice-Test zu den Abbildungen	225
9.2 Darstellung der kognitiven Operationen durch die Materialien	227
9.2.1 Fragebogen zur Einschätzung der Schüler.....	227
9.2.2 Anforderungen durch Text als Input.....	228
9.2.3 Anforderungen durch Experten-Concept-Maps als Input.....	230
9.2.4 Anforderungen des Multiple-Choice-Tests zur Diagnose	231
9.2.5 Anforderungen an Schüler durch den Einsatz von Abbildungen und Modellen	232
9.2.6 Lernaufgaben 1 bis 4 mit kognitiven Anforderungen.....	235
9.2.7 Aufführung der Hilfen zu den Lernaufgaben.....	242
9.2.8 Das Erstellen von gestuften Hilfen nach Stäudel (2008).....	248
9.3 Verwendete Hard- und Software	251
9.4 Ausschreibung der Lehrerfortbildung.....	252
9.5 Genehmigung der Studie.....	253
9.6 Anschreiben an die Lehrkräfte.....	254
9.7 Anschreiben an die Erziehungsberechtigten.....	255
9.8 Augenmodell.....	257
9.9 Weitere Ergebnisse der systematischen Literaturrecherche	258
10 Eidesstattliche Erklärung	261

ABKÜRZUNGEN

Abb.	Abbildung
Abs.	Absatz
APV	Aufgaben-Problem-Verständnis
Art.	Artikel
Aufg.	Aufgabe
β	Teststärke (Beta)
bzw.	beziehungsweise
<i>CI</i>	Konfidenzintervall (confidence interval)
<i>d</i>	Effektstärke (Cohen's <i>d</i>)
<i>df</i>	Freiheitsgrade der Hypothese (degrees of freedom)
ECM	Experten-Concept-Map
η^2	Eta-Quadrat
etc.	und so weiter
<i>F</i>	Prüfgröße der Varianzanalyse
FH	fachliche Hilfe
ff.	und die folgenden (Seiten)
<i>m</i>	Mittelwert (mean)
MC	Multiple-Choice
min	Minimum
max	Maximum
<i>n</i>	Stichprobenumfang
<i>p</i>	Signifikanz
S.	Seite
<i>SD</i>	Standardabweichung (standard deviation)
Tab.	Tabelle
z. B.	zum Beispiel

ABBILDUNGSVERZEICHNIS

Abb. 1: Graphische Darstellung der einzelnen Schritte in zeitlicher Abfolge mittels Gantt-Diagramm.	5
Abb. 2: Ablauf und Ergebnis der systematischen Literaturrecherche.	10
Abb. 3: Inklusion im Vergleich nach Scholz (2012, S.15) in Anlehnung an Sander (2006).	14
Abb. 4: Möglichkeiten von Unterrichtsabläufen auf der Mikroebene (entnommen aus Bohl 2012, Folie 18).	16
Abb. 5: Überblick über Differenzierungsmöglichkeiten in Deutschland (in Anlehnung an Scholz 2012).	18
Abb. 6: Aspekte innerer Differenzierung in Anlehnung an Krüger (2009, Seite 9).	22
Abb. 7: Anforderungen an Lehrkräfte, die durch Binnendifferenzierung entstehen (in Anlehnung an Volkholz et al. 2013, S. 32 f.).	29
Abb. 8: Zusammenstellung wichtiger Aspekte innerer Differenzierung (in Anlehnung an Bohl (2010), Hellrung (2009), Krüger (2009), Leisen (2006), Pietsch (2010), Scholz (2012), Stäudel (2009), Tomlinson (2003), Wodzinski (2007)).	35
Abb. 9: Ebenen externer Repräsentationen im Biologieunterricht in Anlehnung an Leisen (2005, S.8).	36
Abb. 10: Darstellung kognitiver Belastungen durch Unterricht und Lehrmaterialien (in Anlehnung an Sweller et al. 1998).	38
Abb. 11: Kognitive Lerntheorie nach Mayer (2001).	41
Abb. 12: Verarbeitungstiefe basierend auf Horz und Schnotz (2010).	42
Abb. 13: Der Aufgabentyp in Abhängigkeit zu den Grundeinheiten einer Aufgabe (nach Bruder 2003).	60
Abb. 14: Zusammenstellung wesentlicher Aspekte von Aufgaben im Unterricht.	70
Abb. 15: Beschreibungsansätze für den Schwierigkeitsgrad von Aufgaben.	77

Abb. 16: Vorschlag eines Strukturmodells physikalischer Kompetenz als Basis zur gezielten Generierung des Schwierigkeitsgrades bei Aufgaben (entnommen aus Neumann (2007), Seite 111).	82
Abb. 17: Materialentwicklung zu innerer Differenzierung im Rahmen einer Lehrerfortbildung.	91
Abb. 18: Darstellung des symbiotischen Prozesses bei der Entwicklung der Differenzierungsmethode.	100
Abb. 19: Prozess der zyklischen formativen Evaluation.	109
Abb. 20: Die Experten-Concept-Map als fachlicher Input (ohne Farbleitcode).	111
Abb. 21: Beschreibung der Elemente eines Boxplots.	117
Abb. 22: Sprechzeit des Versuchsleiters bei leistungsstärkeren Schülern (n=31) im Vergleich zu leistungsschwächeren Schülern (n=24) als Boxplot (siehe Abb. 21 auf Seite 117).	135
Abb. 23: Sprechzeiten der Lernenden bei leistungsstärkeren Schülern (n=59) im Vergleich zu leistungsschwächeren Schülern (n=49) als Boxplot (siehe Abb. 21 auf Seite 117).	135
Abb. 24: Streudiagramm der Punktezahl im MC-Test und der Zeit zum Ausfüllen des MC-Tests in Sekunden bei allen Schülern.	137
Abb. 25: Streudiagramm der Punktezahl im MC-Test und der Zeit zum Ausfüllen des MC-Tests in Sekunden bei leistungsstärkeren Schülern.	138
Abb. 26: Streudiagramm der Punktezahl im MC-Test und der Zeit zum Ausfüllen des MC-Tests in Sekunden bei leistungsschwächeren Schülern.	138
Abb. 27: Bearbeitungszeit aller Schüler (n=84) bei den Lernaufgaben 1 bis 4 in Sekunden.	139
Abb. 28: Bearbeitungszeit für die Aufgaben 1 bis 4 im Vergleich von leistungsstärkeren zu leistungsschwächeren Schülern in Sekunden.	140
Abb. 29: Bearbeitungszeit für die Aufgaben 1 bis 4 im Vergleich von Schülern mit Text zu Schülern mit Experten-Concept-Map (ECM) in Sekunden.	141

Abb. 30: Vergleich der Bearbeitungszeit von männlichen und weiblichen Schülern bei den Aufgaben 1 bis 4 in Sekunden.....	142
Abb. 31: Bearbeitungszeit für die Aufgaben 1 bis 4 im Vergleich von Schülern mit einem guten Ergebnis zu Schülern mit einem schlechten Ergebnis im Multiple-Choice-Test.	143
Abb. 32: Prozentuale Häufigkeit des Einsatzes der Hilfen bei allen Dyaden. Bei der Beschriftung der x-Achse steht A für Aufgabe und H für Hilfe. Der Balken über A1H1 stellt demnach die prozentuale Häufigkeit von Hilfe 1 bei Aufgabe 1 bei allen Interviews dar (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).....	149
Abb. 33: Vergleich der prozentualen Häufigkeit der Hilfen zwischen leistungsstärkeren und leistungsschwächeren Dyaden. Marginale und signifikante Unterschiede sind durch einen schwarzen Rahmen gekennzeichnet (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).	150
Abb. 34: Vergleich der prozentualen Häufigkeit der Hilfen zwischen Dyaden mit Text und Dyaden mit Experten-Concept-Map. Signifikante Unterschiede sind durch einen schwarzen Rahmen gekennzeichnet (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).....	151
Abb. 35: Vergleich der prozentualen Häufigkeit der Hilfen zwischen Dyaden mit Mädchen und Dyaden mit Jungen. Bei Aufgabe 3 und besonders bei Aufgabe 4 geht die Häufigkeit der beanspruchten Hilfen zwischen den Geschlechtern stark auseinander (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).....	152
Abb. 36: Vergleich der prozentualen Häufigkeit der Hilfen bei Schülern mit einem guten Ergebnis im Multiple-Choice-Test und bei Schülern mit einem schlechten Ergebnis im Multiple-Choice-Test. Signifikante und marginal signifikante Unterschiede zwischen den Gruppen sind mit einem schwarzen Rahmen markiert (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).	153
Abb. 37: Prozentuale Häufigkeit der Qualität der Schülerantworten bei den Aufgaben 1 bis 4.....	162

Abb. 38: Prozentuale Häufigkeit der Qualität der Schülerantworten bei den Aufgaben 1 bis 4 im Vergleich zwischen leistungsstärkeren und leistungsschwächeren Dyaden. Leistungsstärkere Dyaden erzielten bei allen Aufgaben prozentual eine gute Antwort.	163
Abb. 39: Prozentuale Häufigkeit der Qualität der Schülerantworten bei den Aufgaben 1 bis 4 im Vergleich zwischen Dyaden mit Text und Dyaden mit Experten-Concept-Map als Input.....	164
Abb. 40: Prozentuale Häufigkeit der Qualität der Schülerantworten bei geschlechtshomogenen Dyaden.	165
Abb. 41: Prozentuale Häufigkeit der Qualität der Schülerantworten bei Dyaden mit einem guten Ergebnis im Multiple-Choice-Test im Vergleich zu Dyaden mit einem schlechten Ergebnis im Multiple-Choice-Test.....	166
Abb. 42: Ablauf der Interviews mit Aufgabentexten (siehe Seite 112) und Schwierigkeitsgraden (siehe Seite 22) entsprechend der Anforderungsbereiche der KMK (2004) und dem Kompetenzstufenmodell nach Mayer (2013).	171
Abb. 43: Das Entwickeln von geschlechtstypischem Verhalten (Sievert und Kröhnert 2015, S. 31).	182
Abb. 44: Lösungsansätze zum Umgang mit Geschlechterunterschieden im naturwissenschaftlichen Unterricht (nach Sievert und Kröhnert 2015, S. 39).	187
Abb. 45: Ergänzende Abbildungen zum Input für alle Schüler (ohne Farbleitcode).	232
Abb. 46: Faltanweisung der Druckvorlage von Stäudel (2008, Seite 6).	249

TABELLENVERZEICHNIS

Tab. 1: Differenzierungsmöglichkeiten über den Lernweg (entnommen aus Scholz 2008, S. 15).	22
Tab. 2: Aspekte professioneller Kompetenz von Lehrkräften (entnommen aus Volkholz et al. 2013, S. 33).	27
Tab. 3: Anforderungen an Lehrkräfte, die aus Binnendifferenzierung erwachsen (in Anlehnung an Hellrung 2009, S. 113).	28
Tab. 4: Zusammenstellung subjektiver Fehlerquellen bei der Einschätzung von Schülern durch Lehrer. Hierbei handelt es sich um eine Zusammenstellung aus Sicht der Schulpraxis (teilweise nicht empirisch belegt, aber dennoch in der Praxis relevant).	30
Tab. 5: Die Bedeutung von unterschiedlichen Faktoren für erfolgreiches schulisches Lernen: starke und sehr starke positive Effekte (in Anlehnung an Hattie 2009 entnommen aus Köller 2012, S. 3).	33
Tab. 6: Abgrenzung zwischen Mind-Map und Concept-Map (in Anlehnung an Eppler 2006, S.203).	46
Tab. 7: Phasen der Erstellung einer Concept-Map.	49
Tab. 8: Unterrichtsphasen und Aufgabenziele in Anlehnung an Häußler (Häußler und Lind 2000).	65
Tab. 9: Unterschiede zwischen Test- und Lernaufgaben (in Anlehnung an Leisen 2010, S. 11).	67
Tab. 10: Merkmale didaktischer und fachlicher Aufgabenqualität mit Analysekriterien als Erweiterung zu Blömeke (Blömeke et al. 2006, S. 337).	73
Tab. 11: Interessegenerierende Kontexte in Anlehnung an Hammann (Hammann 2006).	74
Tab. 12: Kompetenzmatrix der Bildungsstandards nach Leisen (Leisen 2006, S.261).	81
Tab. 13: Kompetenzstufen im Bereich Fachwissen nach Mayer (Mayer et al. 2013)..	83

Tab. 14: Die dargestellten Kriterien ergeben sich aus den im Text beschriebenen Theorien und sind eine Gegenüberstellung der unterschiedlichen Perspektiven auf den Schwierigkeitsgrad von Aufgaben. Die fett gedruckten Theorien wurden bei der Aufgabenkonstruktion berücksichtigt. Die Klassifikation der Aufgaben dieser Studie befindet sich in Tab. 18 auf Seite 113.....	85
Tab. 15: Merkmale wirksamer Lehrerfortbildungen nach Lipowsky (Lipowsky und Rzejak 2012, S. 5 ff.) und deren Umsetzung in dieser Studie.	90
Tab. 16: Themenschwerpunkte der Lehrerfortbildung.	93
Tab. 17: Der Text als fachlicher Input (ohne Farbleitcode).	110
Tab. 18: Einordnung des Schwierigkeitsgrades der Aufgaben dieser Studie.	113
Tab. 19: Analysemethoden zur Beurteilerübereinstimmung in Abhängigkeit zur Skalierung der Daten (in Anlehnung an Dorsch et al. 2014).	119
Tab. 20: Umsetzung der Kriterien nach Bayerl (2011) zur Einschätzung der Beurteilerübereinstimmung in dieser Studie (Interrater Agreement).	122
Tab. 21: Nichtparametrische Korrelationen nach Kendalls Tau (b) zur Beurteilerübereinstimmung (signifikante Werte sind fett gedruckt).	123
Tab. 22: Zusammenstellung der hier bearbeiteten Forschungsfragen.	125
Tab. 23: Deskriptive Daten zum Multiple-Choice-Ergebnis bei leistungsstärkeren und leistungsschwächeren Schülern in Abhängigkeit vom Input.	128
Tab. 24: Das Ergebnis der zweifaktoriellen Varianzanalyse ohne Messwiederholung (Input und Leistungsgruppe) zeigte einen signifikanten Effekt von Experten-Concept-Maps auf das Multiple-Choice-Ergebnis aller Schüler (signifikante Ergebnisse sind fett gedruckt, interessante Ergebnisse sind fett und kursiv gedruckt).	128
Tab. 25: Das Ergebnis der einfaktoriellen ANOVA zeigte einen marginal signifikanten Effekt von ECMs (ECM = Experten-Concept-Map) auf leistungsschwächer Schüler. Entsprechende Ergebnisse sind fett gedruckt.	129
Tab. 26: Das Ergebnis der zweifaktoriellen ANOVA zeigt einen Haupteffekt des Inputs. Es konnte weder ein Haupteffekt durch das Geschlecht, noch ein Interaktionseffekt zwischen Geschlecht und Input gefunden werden. Dies bedeutet,	

dass sich der Input auf beide Geschlechter positiv auswirkt. Signifikante Werte sind fett gedruckt.	130
Tab. 27: Vergleich der Schulnoten in Bezug auf das Geschlecht der Schüler. Sowohl in Biologie als auch in Deutsch bestehen signifikante Notenunterschiede zwischen den Geschlechtern, außer in Mathematik. Signifikante Werte (p) sind fett gedruckt, sowie deren Effektstärken (d) aufgeführt.	131
Tab. 28: Zeitangaben zu der Dauer der Aufgaben 1 bis 4 in Sekunden (gerundet). ...	133
Tab. 29: Sprechzeit des Versuchsleiters in den Interviews (n=56) und die Sprechzeiten der Schüler (n=108) bei Aufgabe 4 in Sekunden.	134
Tab. 30: Vergleich der Sprechzeiten des Versuchsleiters und der Schüler in Sekunden (Aufgabe 4).	134
Tab. 31: Ergebnisse der linearen Regression in Bezug auf den Vorhersagewert der Schülerleistung im Multiple-Choice-Test in Abhängigkeit zur Bearbeitungszeit. Signifikante Ergebnisse sind fett gedruckt.	137
Tab. 32: Gesamte Anzahl aller Hilfen pro Aufgabe.	144
Tab. 33: Häufigkeit der gegebenen Hilfen pro Aufgabe. Die am häufigsten eingesetzten Hilfen pro Aufgabe sind fett gedruckt.	145
Tab. 34: Anzahl und Aufteilung der Hilfen in zwei Kategorien: Aufgaben-Problem-Verständnis (APV) und rein fachliche Hilfe (FH). Hilfen, die sich auf APV und FH beziehen, sind grau hinterlegt.	145
Tab. 35: In dieser Tabelle wird in jeder Zelle für jede Hilfe zuerst deren Häufigkeit in den Interviews aufgeführt. Die anschließende Abkürzung gibt Auskunft über die Hauptkategorie der Hilfe (APV = Aufgaben-Problem-Verständnis; FH = fachliche Hilfe). Der darauffolgende Satz beschreibt die jeweilige Subkategorie der Hilfe. Die letzte Zahl gibt an, wie viel Prozent der Schüler diese Hilfe im Interview beansprucht haben.	147
Tab. 36: Häufigkeit der Einschätzungen aller Schüler (n=110) zum Schwierigkeits- und Interessantheitsgrad der Aufgaben. Die Schüler nannten jeweils die einfachste, schwerste und interessanteste Aufgabe. Doppelnennungen und Enthaltungen waren erlaubt.	155

Tab. 37: Häufigkeit der Schülereinschätzungen zum Schwierigkeits- und Interessantheitsgrad der Aufgaben in Bezug auf das Geschlecht (♂ n=49 und ♀ n=61). Starke Unterschiede zwischen den Geschlechtern sind fett gedruckt.	155
Tab. 38: Ranking der Schüleraussagen zu den Aufgaben.	160
Tab. 39: Punkteanzahl im Multiple-Choice-Test zu den Abbildungen.	167
Tab. 40: Korrelationen zwischen dem Wissenstest und der Qualität der Schülerantworten für die Aufgaben 1 bis 4 bei unterschiedlichen Schülergruppen. Bei keiner Schülergruppe traten signifikante Korrelationen nach Pearson zwischen den Multiple-Choice-Ergebnissen und der Qualität der Schülerantworten bei den Lernaufgaben auf.	170
Tab. 41: Ablauf der Interviews mit Einsatz der Materialien.	220
Tab. 42: Gegenüberstellung der Anforderungen an Schüler durch Text und Experten-Concept-Map (ECM).	231

1 EINLEITUNG UND ZIELSETZUNG DER ARBEIT

Innere Differenzierung ist ein Überbegriff für wichtige Maßnahmen zur Anpassung von Unterricht an das Niveau von Schülern innerhalb einer Klasse. Aus pädagogischer und gesellschaftlicher Sicht nimmt innere Differenzierung eine entscheidende Rolle im Bildungsprozess ein (Thonhauser 2008, Hänze 2009, Tillmann 2004). Deshalb muss die Anpassung an die individuellen Anforderungen unterschiedlicher Zielgruppen einen zentralen Bestandteil bei der Planung und Durchführung von Unterricht einnehmen, jedoch weisen diese in der Praxis Mängel auf (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung 1997, Bruder 2000, Blume und Rademann 2000, Gogolin et al. 2003, Kunze 2010).

Diese Mängel sind meist an einer nicht systematischen Anwendung von didaktisch-methodischen Prinzipien zu erkennen und spiegeln sich im teilweise sehr willkürlichen Einsatz von Medien und anderen Hilfsmitteln wider (Winter 2008, Gräsel et al. 2004). Unterstützung bedarf es besonders in Hinblick auf die Umsetzung von Differenzierungsmaßnahmen in der Klasse als auch bei der Erstellung von neuen differenzierenden Lehr-Lern-Materialien (Bönsch 2012).

Auch bei der Durchführung von Differenzierungen sind Qualitätsunterschiede zum Teil erheblich und alles andere als zweckdienlich (Krammer 2009). Dies zeigt sich unter anderem am hohen Anteil von Frontalunterricht und wenig interaktiven Methoden und praktischen Übungen (Leisen 2005, Jatzwauk 2007). Um die Qualität von Unterrichtsmaßnahmen zur inneren Differenzierung zu optimieren, liefert die Fachdidaktik Planungs- und Umsetzungskriterien (Leisen 2006, Bräu 2007, Wodzinski und Wodzinski 2007, Krüger und Meyfarth 2009). Darüber hinaus gibt sie Ansätze zur Erstellung neuer Unterrichtsmaterialien (Aufschnaiter 2000, Forschergruppe-Universität Kassel et al. 2007).

Bezüglich der Wirksamkeit gezielter Differenzierung im Biologieunterricht existieren bisher keine ausreichenden empirischen Befunde (Brünken et al. 2005, Nesbit und Adesope 2006, Schneider und Ludwig 2012). Die Evaluation von entwickelten Differenzierungsmaßnahmen ist jedoch äußerst wichtig, da sich ein unbedarfter Einsatz auch kontraproduktiv auswirken kann.

So konnte Kalyuga zum Beispiel zeigen, dass der Einsatz einer Differenzierungsmaßnahme bei leistungsschwächeren Schülern zwar Erfolg hat, bei leistungsstärkeren Schülern jedoch durch den *expertise reversal effect* zu verminderter Lernleistung führt (Kalyuga et al. 2003, Kalyuga 2013, Blayney et al. 2015).

Ziel dieser Arbeit ist es, den Mangel an Differenzierung im Biologieunterricht zu ergründen, einen praxistauglichen Lösungsansatz zu entwickeln und dessen Wirksamkeit empirisch zu überprüfen.

2 PROBLEMSTELLUNG UND AUFBAU DER ARBEIT

„Heterogenität ist keine pädagogische Größe, sondern eine Tatsache“ (Leisen 2006, S. 265). Diese Aussage ist allgemeingültig und von großer Bedeutung. Sie liefert eine Begründung für den offensichtlichen Mangel der Binnendifferenzierung im Biologieunterricht (Jatzwauk und Sandmann 2007, S. 173) und dem daraus resultierenden Bedarf an praxistauglichen Lösungsansätzen sowie einer Überprüfung von deren Wirksamkeit.

Die folgenden Fragestellungen dienen einzig und alleine diesem Zweck und bilden das Grundgerüst dieser Arbeit:

- a) Welche Ansätze innerer Differenzierung eignen sich aus fachdidaktischer Sicht für die Vermittlung biologischer Inhalte? Welche dieser Ansätze nehmen Lehrkräfte aufgrund ihres Erfahrungswissens an?
- b) Hat eine differenzierende Darstellungsform annähernd identischer Lehrinhalte über Experten-Concept-Maps oder Text Auswirkungen auf das Wissen und Können bei Lernenden mit unterschiedlichen Leistungsniveaus?
- c) Treten bei den eingesetzten Differenzierungsmaßnahmen geschlechtsspezifische Unterschiede auf?
- d) Welchen Einfluss hat die Lernzeit auf das Ergebnis der Schüler?
- e) Wie gestaltet sich der Lernprozess der Schüler mit den differenzierten Darstellungsformen und der Lernaufgaben und wo treten dabei Schwierigkeiten auf?

Die Struktur der vorliegenden Arbeit orientiert sich an der Reihenfolge der eben genannten Forschungsfragen. Im folgenden Kapitel werden zunächst verschiedene Theorien relevanter Themen betrachtet.

Anschließend wird Fragstellung a) beantwortet. Hierfür wurden Lehrkräfte zu einer Fortbildung eingeladen. Bei der Durchführung kamen explorative Forschungsmethoden wie Befragung und teilnehmende Beobachtung zum Einsatz. Die Beschreibung der Teilnehmenden, den Ablauf der Lehrerfortbildung und deren Ergebnisse werden in Kapitel 4 ab Seite 89 dargestellt.

Basierend auf der Zusammenarbeit mit den Lehrern wurde gemeinsam eine konkrete Methode zur inneren Differenzierung im Biologieunterricht ausgewählt und für die wissenschaftliche Überprüfung entwickelt. Die Beschreibung dieses Ansatzes sowie die allgemeinen methodischen Vorgehensweisen befinden sich in Kapitel 5 ab Seite 105.

Nach der kooperativen Entwicklung und formativen Evaluation der Materialien wurde deren Wirkung summativ überprüft. Die Ergebnisse zu den Forschungsfragen b), c), d) und e) finden sich in Kapitel 6 ab Seite 125.

In Kapitel 7 werden die Ergebnisse der Fragenblöcke a) bis e) dieser Studie aufgegriffen und diskutiert. Sofern sich aus der Diskussion Schlussfolgerungen für die Schulpraxis im Biologieunterricht ableiten lassen, werden diese zusätzlich zu den jeweiligen Punkten formuliert und im Text hervorgehoben.

Im abschließenden Kapitel befinden sich die ausführlichen Beschreibungen der eingesetzten Materialien. Kapitel 9 beinhaltet darüber hinaus die Nennung und Beschreibung aller anderen Materialien wie zum Beispiel das Anschreiben an die beteiligten Lehrkräfte, die Genehmigung der Studie, die verwendete Hard- und Software.

Auf der folgenden Seite wird der zeitliche Ablauf der Studie in Abb. 1 dargestellt.

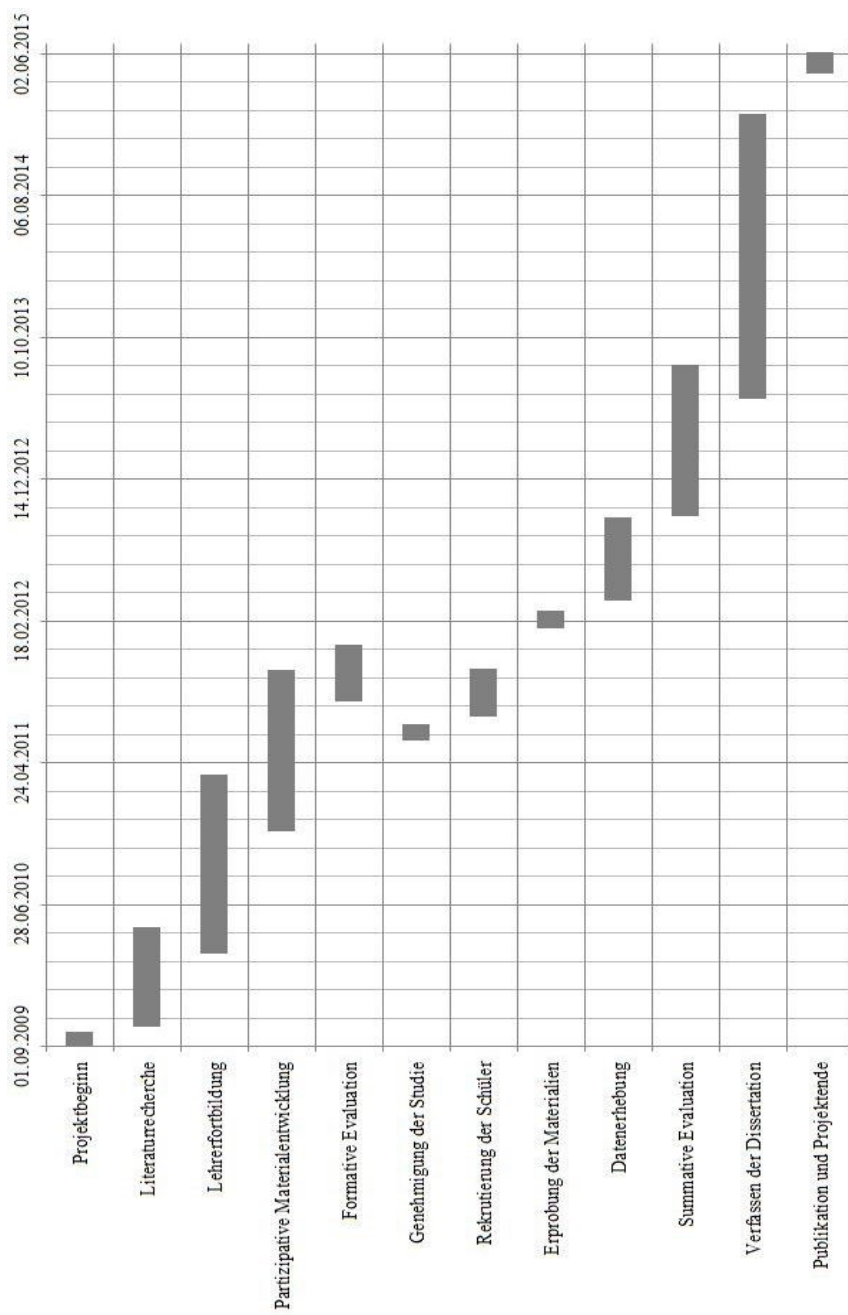


Abb. 1: Graphische Darstellung der einzelnen Schritte in zeitlicher Abfolge mittels Gantt-Diagramm.

3 THEORETISCHE FUNDIERUNG RELEVANTER THEMEN UND INHALTE

Die konsequente Beteiligung von Lehrkräften führte zu einer starken Prozess- und Anwendungsorientierung dieser Arbeit. Die gewollt reflexive Auseinandersetzung zwischen didaktischer Theorie und pädagogischer Praxis erzeugt jedoch ein Spannungsverhältnis. Die Spannung entsteht daraus, dass in der Wirklichkeit einer pädagogisch handelnden Person äußerst selten eine Situation entsteht, in welcher nur eine Theorie Anwendung findet. Die alltägliche Einzigartigkeit des eigenen pädagogischen Handelns ist dabei unter anderem auf der Lehrerpersönlichkeit, der Zusammensetzung der Klasse, der Spontanität einer auftretenden Situation im Unterricht, der individuellen Voraussetzungen der Schule oder des Kollegiums begründet. Die berufsbedingte Variabilität macht die Formulierung von Wenn-Dann-Beziehungen in der Didaktik äußerst schwierig. Im Unterricht kann es dadurch bei der Anwendung von didaktischen Theorien zu Frustration bei Praktikern kommen, denn didaktische Theorie ist in der Praxis in der Regel nicht 1:1 anwendbar. Sie muss immer einen Handlungsspielraum lassen und erfordert dadurch meistens eine Anpassung an die jeweilige Situation. Aus diesem Grund können aus didaktischer Theorie selten genaue Handlungsanweisungen abgeleitet werden. Daraus folgt, dass es nicht für jede Situation im Unterricht eine passende Theorie geben kann. Zum anderen folgt dadurch aber auch, dass man sich bei der Planung und Konzeption von Handlungsanweisungen auf allgemein anwendbare Theorien wie die Folgenden berufen muss:

- Allgemeine Aspekte von Binnendifferenzierung in Schule und Unterricht, Voraussetzungen für das Gelingen, Gütekriterien, sowie damit einhergehende Veränderungen der Lehrerrolle.
- Lehren und Lernen mit Repräsentationen (Medien), ausgehend von der begrenzten Lernleistung der Cognitive Load Theory über allgemein gültige Prinzipien zu Medien zu speziellen Gütekriterien des Einsatzes unterschiedlicher Lehrmaterialien im Verbund (multi media) mit explizitem Bezug auf Individualisierung und deren Konsequenzen auf den Biologieunterricht.

In Anlehnung an das Thema Binnendifferenzierung in der Schule und den dazugehörigen Erkenntnissen aus der Zusammenarbeit mit unseren Lehrkräften (siehe Kapitel 4.3 Seite 97 ff.), werden in den folgenden Kapiteln die relevanten allgemein anwendbaren Theorien in den jeweiligen Disziplinen vorgestellt. Diese beinhalten:

- Vorstellung der Methode des Concept-Mapping und Experten-Concept-Maps, gefolgt von der Aufführung von Qualitätskriterien in Bezug auf Lernzuwachs und Individualisierung.
- Darstellung formaler Aspekte und Funktionen von Lernaufgaben im Unterricht, einer Abgrenzung zu Testaufgaben gefolgt von Qualitätskriterien guter Aufgaben sowie Einstufungsmöglichkeiten des Schwierigkeitsgrades von Aufgaben im Biologieunterricht.

Die fachwissenschaftlichen Inhalte der Aufgaben beziehen sich auf das Niveau des Mittleren Schulabschlusses in Biologie für die Realschule in Baden-Württemberg. Relevante Inhalte entsprechen daher dem Basiswissen der Biologie im Bereich der Sinneswahrnehmungen. Grundlage für die Konstruktion aller Aufgaben und Materialien sind Lehrwerke der Biologie (siehe Campbell et al. 2009, Markl und Sadava 2011, Schmidt et al. 2000).

Für die Konzeption und Durchführung der Lehrerfortbildung sowie für die Konstruktion von Aufgaben zur individuellen Förderung ist eine theorie- und evidenzbasierte Fundierung folgender Aspekte von Bedeutung: Differenzierung im Unterricht sowie Lehren und Lernen mit Repräsentationen und Lernaufgaben. Zur theoretischen Rahmung der genannten Themen wurde eine allgemeine Literaturrecherche durchgeführt, auf welche im Folgenden eingegangen wird.

3.1 Methodisches Vorgehen

Allgemeine Literaturrecherche

Die theoretische Fundierung der Themenblöcke: Differenzierung im Unterricht, Lehren und Lernen mit Repräsentationen sowie Aufgaben im Unterricht basieren primär auf aktuellen, international anerkannten Lehrwerken. Entsprechend der Ausführlichkeit der Lehrwerke wurden diese gegebenenfalls durch weiterführende und tiefergehende Literatur ergänzt.

Die Ergänzung der Lehrwerke beinhaltete sowohl das Screening der jeweiligen Literaturverzeichnisse der Lehrwerke als auch eine Recherche in den gleichen Datenbanken wie bei der systematischen Literaturrecherche mit folgenden Schlagworten:

- Aufgabenorientierung, Schulaufgaben, Lernaufgaben, Lerndifferenzierung, Leistungsdifferenzierung, Unterschied, Heterogenität und Leistungsunterschiede, Aufgabenpräsentation, Aufgabenkultur, Leistung*, Lern*, Schulaufgab*, Unterschied,

Lernen, Heterogenität, Kompetenz*, Aufgabenanforderung, Anforderung*, differenz*, Aufgabenstellung, Klassenkomposition, Aufgabenstruktur, Aufgabenfunktion und Komplexität, Kompositionseffekt, Leistungsbeurteilung, task based learning und task orientation, task effect, task demand, complexity, assessment*, learning*, task*, school*, task composition, task demand*, task presentation*, task complexit*, differential*, difference*, differen*, task effect*, heterogeneity, task instruction*.

Über den Zeitraum der systematischen Literaturrecherche hinaus wurde bis zur Einreichung der Arbeit weiterhin Literatur per Handsuche gesichtet und hinzugefügt.

Systematische Literaturrecherche zur Binnendifferenzierung

Da der Fokus auf dem Umgang mit Heterogenität im Biologieunterricht liegt, wurde eine systematische Literaturrecherche ausschließlich in Bezug auf innere Differenzierung im Biologieunterricht durchgeführt.

Um den Forschungsstand zur inneren Differenzierung im Biologieunterricht zu erheben, wurde von Oktober 2009 bis April 2010 in den Datenbanken FIS Bildung, ERIC, PSYINDEX, PEDOCS und EBSCO recherchiert.

Hierbei wurden die Datenbanken mit folgenden Schlagwörtern in der Freitextsuche durchsucht: Binnendifferenzierung, innere Differenzierung, Heterogenität und Biologie. Zusätzlich wurde die Suche mit den entsprechenden englischen Begriffen ergänzt. Die Suche erfolgte unter Verwendung von trunkierten¹ Begriffen und Booleanen² (zum Beispiel *internal differen** OR *inner differen** OR *heterog** AND *biolog**). Darüber hinaus erfolgte eine Handsuche in der Zeitschrift *Didaktik der Naturwissenschaften* (ZfDN) aus den Jahren 2000 bis 2012. Daraufhin wurden die Referenzlisten aller eingeschlossenen Publikationen nach weiteren relevanten Arbeiten gescreent. Das Ergebnis der systematischen Literaturrecherche erbrachte insgesamt 310 Treffer.

Im Anschluss daran fand eine erste Auswahl statt. Hierbei wurden alle Treffer ausgeschlossen, die nicht auf Deutsch oder Englisch verfasst waren und vor dem Jahr 2000 publiziert wurden. Hiermit reduzierte sich die Anzahl der Treffer auf 85. Daraufhin wurden die Titel und deren Abstracts gelesen und danach aussortiert, ob es sich bei diesen Arbeiten tatsächlich um Differenzierung im Unterricht handelt und nicht um

¹ Das *-Symbol steht hierbei für einen oder mehrere undefinierte Charaktere im Suchbegriff, die sogenannte Wildcard, welche durch *Trunkierung* eine erweiterte Suche in Datenbanken erlaubt.

² Ein *Boolean* erlaubt die Verknüpfung *AND/OR* oder den Ausschluss *NOT* von Suchbegriffen bei der Datenbankrecherche, was zu einer gezielten Eingrenzung der Ergebnisse führt.

bildungspolitische Vergleichsstudien, Bildungsdiskussionen, Handreichungen für den Unterricht oder politische Forschungen. Am Ende blieben insgesamt 7 Treffer übrig. Abschließend wurde überprüft, ob es sich hierbei um theoretische oder empirische Beiträge zu Binnendifferenzierung und Leistungsheterogenität im Biologieunterricht handelte.

Das Ergebnis brachte keine Treffer für empirische Befunde zum differenzierten Input bei Lernenden mit unterschiedlichen Leistungsniveaus im Fach Biologie. Gefunden wurden dafür aber eine empirische Arbeit zum Vorkommen von Differenzierung im Unterricht (Jatzwauk et al. 2008). Hierbei handelt es sich um Videoanalysen von Biologieunterricht, welche zeigen, dass Binnendifferenzierung hier bisher kaum stattfindet. Darüber hinaus finden sich Arbeiten, die im Bereich der fachdidaktischen Entwicklungsforschung zu verorten sind und allgemeine theoretische Ansätze zu Differenzierung liefern (Leisen 2006, Bräu 2007, Krüger und Meyfarth 2009). Der Ablauf der systematischen Literaturrecherche mit genauen Trefferzahlen ist in der folgenden Abbildung dargestellt. Ungefilterte Titel zu Differenzierung mit explizitem Bezug auf Biologieunterricht befinden sich im Anhang ab Seite 258.

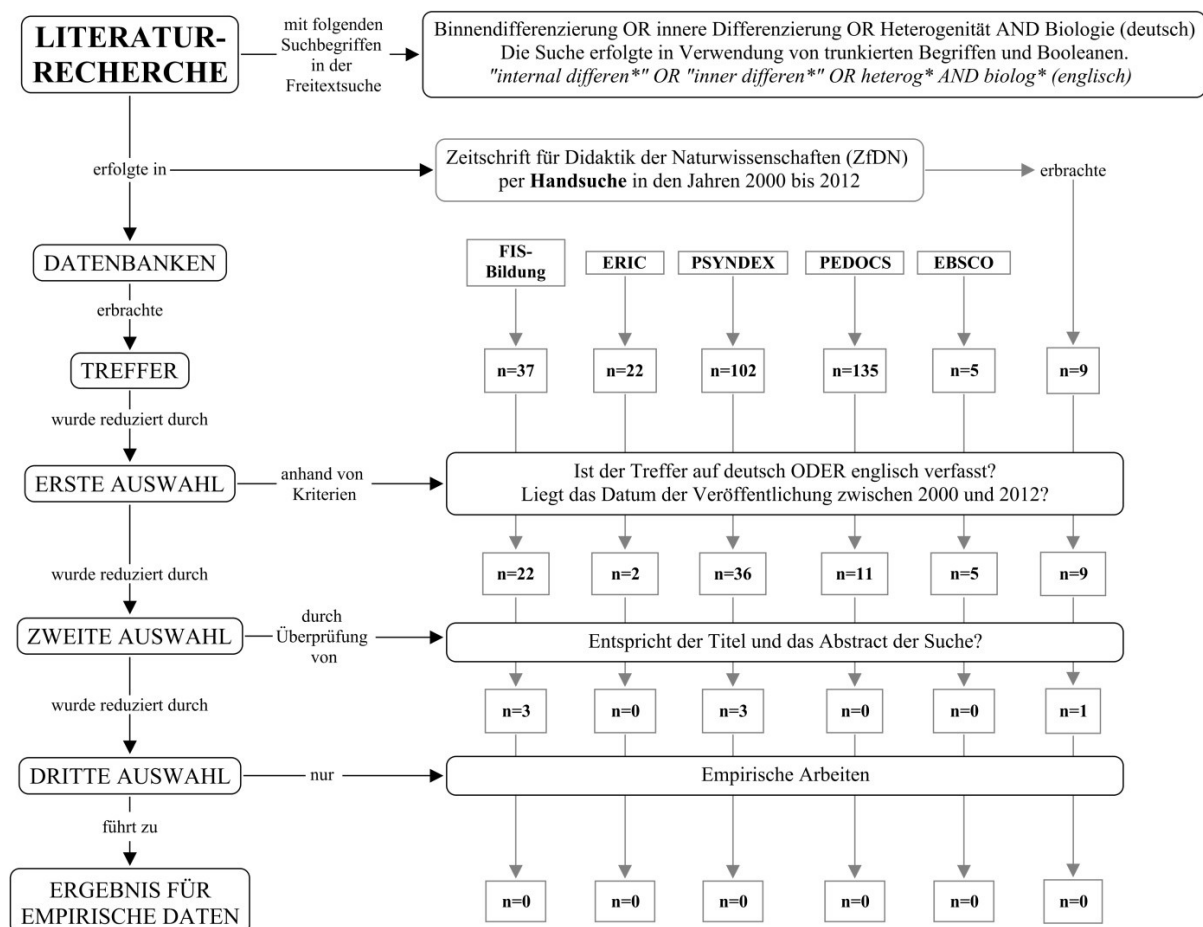


Abb. 2: Ablauf und Ergebnis der systematischen Literaturrecherche.

3.2 Differenzierung im Unterricht

Ausgehend von dem Großthema der Binnendifferenzierung werden in den folgenden Abschnitten zunächst die gesellschaftlichen Diskussionen über Differenzierung in Deutschland dargestellt. Anschließend wird die Anwendung von Differenzierung theoretisch begründet und in direkten Bezug zur Schule gesetzt. Nach der Nennung allgemeiner Aspekte von Differenzierung auf Schulebene wird der thematische Rahmen der Arbeit gesetzt. Nach einer Trennung zwischen äußerer und innerer Differenzierung wird im letzten Drittel ausführlich auf innere Differenzierung im Unterricht eingegangen. Dies beinhaltet sowohl die Nennung von Voraussetzungen für das Gelingen von Differenzierung im Unterricht, als auch das Aufzählen von methodischen Möglichkeiten und Unterrichtsabläufen. Daraufhin werden auf Vor- und Nachteile von Diagnosen im schulischen Kontext eingegangen und abschließend allgemeine Prinzipien von Differenzierung im Unterricht aufgeführt.

Die Differenzierungsdebatte seit dem Jahr 2000

Die Diskussion über Differenzierung ist nicht neu und wird seit der Zeit der Reformpädagogik vehement geführt. Heutzutage ist das nicht anders als damals, allerdings erlebt die Diskussion eine neue Blüte. Konzepte der inneren Differenzierung erfreuen sich größter Beliebtheit in Lehrerhandreichungen und Methodenratgebern, wohingegen sie in der Forschung eher ein Schattendasein führen (Bohl et al. 2012, S. 6). Die Popularität ist nicht verwunderlich, denn Lehrkräfte sollen heutzutage das Kunststück fertigbringen, die Lernprobleme und Schwierigkeiten ihrer Schüler nebeneinander und am besten noch gleichzeitig zu überwinden (Becker 2009, S. 7). Aus der Notlage heraus werden Themenhefte und Sammelwerke erstellt. Beim Vergleich der Handreichungen kommt allerdings heraus, dass in über 20 Werken Heterogenität zwar als Chance begriffen wird, jedoch keiner sagt, worin diese Chance genau besteht (Budde 2012, S. 15).

Im Vergleich der damaligen Differenzierungsdiskussionen zu heute lassen sich jedoch Unterschiede finden. Mittlerweile geht es bei den Debatten hauptsächlich um die Individualisierung von Lernprozessen und kooperative Lernformen. Hierbei wird vor allem die pluralistische Gesellschaft betont, in der Mehrsprachigkeit und Multiperspektivität produktiv genutzt werden sollte, um unterschiedliche Lösungen für Probleme zu erarbeiten, Informationen zu kommunizieren und Werte und Interessen zu diskutieren (Walgenbach 2013, S. 7). Aus schulischer Sicht ist dies unbestritten, allerdings darf man sich von den neuen, fortschrittlichen und teilweise sogar revolutionären Hilfen und Handreichungen für den modernen Unterricht nicht blenden lassen. Nach Helmke (Helmke 2006b, S. 1–2) ist die Gleichsetzung von *innovativ* mit *gut* nämlich

genauso naiv wie falsch, denn je nach Unterrichtsqualität gibt es den brillanten und anregenden lehrerzentrierten Frontalunterricht genauso wie den dilettantischen Gruppen- oder Projektunterricht. Die Diskussion über eine neue Unterrichtskultur entspricht daher nicht der Abschaffung der älteren und bewährten Lehrformen, sondern einer Erweiterung des Unterrichtsspektrums, die geeigneter ist, die Entwicklung von methodischen, personalen und sozialen Kompetenzen bei Schülern zu fördern (Scholz 2012, S. 20).

Begründung von Differenzierung auf institutioneller Ebene

Zur Begründung von Differenzierung in der Schule lassen sich einige Punkte aufzählen. Zu den wenig diskutierbaren zählen Gesetze. Entsprechend des Geltungsbereichs und der Tragweite eines Gesetzes lässt sich Differenzierung aus geltendem Recht von New York bis Stuttgart ableiten. So stimmte der Bundestag Ende 2008 dem Artikel 24 der Vereinten Nationen für Rechte von Menschen mit Behinderung zu. Diese UN-Konvention soll das Recht auf Chancengleichheit für alle Menschen in den jeweiligen Bildungssystemen ohne Diskriminierung gewährleisten (Bundestag 21.12.2008, S. 1436). Differenzierung lässt sich nicht nur international sondern auch national juristisch begründen. In Deutschland steht das gesamte Schulwesen unter Aufsicht des Staates (Art. 6 Grundgesetz). Bildung hingegen ist Sache der Bundesländer. Bei einem Blick in die Landesverfassung und das Schulgesetz von Baden-Württemberg steht schon in § 1, Abs. 1, dass „jeder junge Mensch ohne Rücksicht auf Herkunft oder wirtschaftliche Lage das Recht auf eine seiner Begabung entsprechende Erziehung und Ausbildung hat“. Ferner steht in Abs. 2 des gleichen Paragraphen, dass die Schule dazu angehalten ist, die Schüler in der Entfaltung ihrer Persönlichkeit und Begabung zu fördern und dass das Schulwesen nach diesem Grundsatz gestaltet werden soll (Verfassung des Landes Baden-Württemberg vom 11. November 1953 und Schulgesetz, vom 01.08.1983).

Diskrepanz zwischen gesetzlichen Vorgaben und schulischer Realität

Vergleicht man die gesetzlich geforderten Unterstützungen und die damit verbundenen Rahmenbedingungen mit der Realität in der Schule, so weichen diese stark voneinander ab. PISA bescheinigt dem deutschen Bildungssystem im internationalen Vergleich inzwischen mehrfach, eine Art negatives Musterland zu sein (Leisen 2006, S. 265). Bei den Vergleichen der Bildungssysteme zeigt sich immer wieder, dass die mit Migrationshintergrund verbundenen Lebensbedingungen im Vergleich immer höher sind als in den meisten anderen OECD-Mitgliedsstaaten (Stanat et al. 2010, S. 200). Bei einem Blick in die Klassen und den laufenden Unterricht scheint das nicht verwunderlich zu sein. In Deutschland orientiert sich der Unterricht nach Tillmann an den *Mittelköpfen*

(dem Durchschnitt), wodurch alle Schüler außerhalb des Mittelbereichs Gefahr laufen, zu Problemfällen zu werden (Tillmann 2004, S. 7). Scholz nennt dies das *fiktive Mittelmaß* und beschreibt standardisierten Normalunterricht mit den 7-G-Prinzipien: „Die gleichen Schüler lösen beim gleichen Lehrer im gleichen Raum zur gleichen Zeit im gleichen Tempo die gleichen Aufgaben mit dem gleichen Ergebnis“ (Scholz 2012, S. 12). Die etwas provokante Beschreibung von Unterricht macht allerdings deutlich, dass standardisierter Unterricht nicht möglich ist, da Unterricht immer in irgendeiner Form angepasst werden muss. Denn insbesondere schwächere Lernende tendieren schnell dazu, an Ausdauer und Interesse zu verlieren, wenn eine Aufgabe nicht zu ihren Kompetenzen passt (Leisen 2006, S. 262). Das Scheitern von Schülern führt dabei im besten Fall zu Langeweile und im schlimmsten Fall zu Frustration. Beides führt über kurz oder lang jedoch zu Unterrichtsstörungen, wodurch wiederum die gesamte Klasse beeinträchtigt wird.

Anstatt den Unterricht nach einem fiktiven Mittelmaß auszurichten, sollte man der Vielfalt der Schüler mit der Vielfalt der Lehre begegnen und durch differenzierende Maßnahmen auf die Bedürfnisse möglichst aller Lernenden eingehen (Bohl et al. 2012, S. 6, Scholz 2012, S. 13). Dieser bewusste Umgang mit Heterogenität entspricht nicht nur den Prinzipien für guten Unterricht (Helmke 2006b, S. 5), sondern ist nachweislich ein zentraler Bestandteil für die Verbesserung von Schulleistungen (Klieme et al. 2010, S. 291).

Es sollte nicht nur allgemein sondern auch fachlich differenziert werden

Neben den allgemeindidaktischen und pädagogischen Argumenten gibt es auch noch fachliche Gründe für die Anpassung von Unterricht. So waren in allen PISA-Studien neben der Lesekompetenz besonders mathematische und naturwissenschaftliche Kompetenzen bei Schülern mit Migrationshintergrund deutlich geringer als bei Schülern ohne Migrationshintergrund (Stanat et al. 2010). Von daher müssen Bemühungen zur Verbesserung der schulischen Situation von Heranwachsenden mit Migrationshintergrund unternommen werden, um das sprachliche und fachliche Lernen zu verbessern (Stanat et al. 2010, S. 202). Abgesehen von den nötigen sprachlichen Unterstützungen bestehen in den Naturwissenschaften noch ganz andere Unterschiede zwischen Lernenden, die nicht mit der Herkunft zusammenhängen. So lag der Anteil der Schüler, deren mathematische und naturwissenschaftliche Fähigkeit am Ende der achten Klasse dem Niveau eines erweiterten Grundschulwissens entsprach, bei etwa 20% (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung 1997, S. 67, Fischer et al. 2003, S. 180).

3.2.1 Allgemeine Aspekte von Differenzierung in der Schule

Das Wort *Differenzierung* wird schwammig und diffus verwendet, denn es fehlt an theoretischer Fundierung und Präzisierung dieses Begriffs (Bohl et al. 2012, S. 6). In Anbetracht der bereits erwähnten hitzigen Debatten über das Thema wären manche Menschen geradezu erstaunt, wie viel Differenzierung das deutsche Schulsystem bereits zu bieten hat (siehe Abb. 5 auf Seite 18). Daher erfolgt eine Klärung der Begrifflichkeiten.

Differenzierung und Inklusion

Der Begriff der Differenzierung sollte im öffentlichen Diskurs nicht als Standardantwort auf Probleme in Gesellschaft und Schule verwendet, sondern eher als Sammelbegriff für Möglichkeiten verstanden werden (Wischer 2012, S. 36).

Genauso diffus wird das Wort *Inklusion* gebraucht. Der Begriff Inklusion wird annähernd genauso häufig verwendet wie Differenzierung, manchmal in einem ähnlichen Zusammenhang, manchmal sogar als Synonym. Das ist irreführend. Die vergleichende Darstellung in Abb. 3 verdeutlicht das Prinzip der Inklusion sehr anschaulich.

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 3: Inklusion im Vergleich nach Scholz (2012, S.15) in Anlehnung an Sander (2006).

Exklusion beschreibt in Bezug auf Menschen die Ausgrenzung von Personen aus einer Gruppe aufgrund eines Merkmals. Separation bezeichnet die Trennung von Personen durch ein Merkmal in zwei Gruppen, welche räumlich voneinander getrennt sind und das auch bleiben. Integration ist die Eingliederung von Personen mit einem Merkmal, welches dafür sorgt, dass diese innerhalb der Gruppe anders behandelt werden. Inklusion beschreibt die Zusammensetzung einer Gruppe unter Ausschluss von Merkmalen. Rein hypothetisch lässt sich ein Szenario für jede dieser vier Formen vom Umgang mit Menschen konstruieren, das eine Trennung oder Zusammenführung von Personen rechtfertigen würde (Hinz 2003). Ob diese erdachten Szenarien gut oder schlecht sind, hängt vor allem von dem Merkmal ab, nach dem Menschen selektiert werden. In der momentanen Diskussion über Inklusion bezieht sich das Unterscheidungsmerkmal primär auf Menschen mit Migrationshintergrund und Menschen mit einer Behinderung. Diese Arbeit bezieht sich auf keines der genannten Merkmale, sondern auf Leistungsunterschiede. Des Weiteren geht es bei der Inklusionsdebatte eher um soziale

Entwicklungen und bildungspolitische Entscheidungen als um konkreten Unterricht. Der in dieser Arbeit verwendete Differenzierungsbegriff hingegen bezieht sich ausschließlich auf Schule und ganz gezielt auf Unterricht.

Differenzierungsmöglichkeiten auf schulischer Ebene

Auf schulischer Ebene müssen zwei Formen von Differenzierung unterschieden werden: die äußere und die innere Differenzierung. Bei der äußeren Differenzierung werden durch Auswahlverfahren möglichst homogene Gruppen gebildet, welche dann getrennt voneinander unterrichtet werden. Dies geschieht über verschiedene Schulformen, Schulprofile und Jahrgangsklassen. Die äußere Differenzierung wird dabei nochmals in inter- und intraschulische Differenzierung unterteilt (siehe Abb. 5 auf Seite 18).

Das Ziel der homogenen Schülergruppe pro Klasse erweist sich in der Praxis allerdings als Illusion, da Homogenität meistens nur für ein Kriterium hergestellt werden kann (Stäudel 2009b, S. 10, Scholz 2012, S. 14). Nach Scholz können demnach trotz äußerer Differenzierung enorme Unterschiede innerhalb einer Klasse bestehen. So kann ein Schüler beispielsweise im Hinblick auf die sprachliche Kompetenz viel weiter entwickelt sein als im Hinblick auf sein logisches Denkvermögen (intraindividuelle Unterschiede). Die intraindividuellen Unterschiede eines einzelnen Schülers, können natürlich auch zwischen zwei Schülern einer Klasse bestehen (interindividuelle Unterschiede). Die interindividuellen Unterschiede zwischen zwei Schülern können in Bezug auf deren Entwicklungsstand mehrere Jahre betragen. Beim Vergleich aller Schüler einer Klasse lassen sich nach Scholz zwei weitere Ebenen von Heterogenität beschreiben. Zum einen wäre da die vertikale Heterogenität. Sie bezeichnet das unterschiedliche Leistungsvermögen der Schüler, welches sich in vielen Arbeitsbereichen des Unterrichts zeigt, sobald der Anspruch oder die Anzahl der Anforderungen steigt. Und zum anderen die horizontale Heterogenität. Sie zeigt sich in den unterschiedlichen Interessen zu einem Thema, den anderen Lernwegen und abweichenden Zugangs- und Herangehensweisen an Aufgaben (Scholz 2012, S. 9). Neben den leistungsbezogenen Unterschieden innerhalb einer Klasse zählt Scholz weitere Aspekte auf, welche bei der Planung von Unterricht eine Rolle spielen. Diese wären vielfältige kulturelle und nationale Identitäten und den damit verbundenen religiösen Sozialisationen, unterschiedliche familiäre und sozioökonomische Kontexte, abweichende Lern- und Arbeitsverhalten in Bezug auf Zeitmanagement und Arbeitsorganisation, unterschiedliche Leistungsmotivationen auch in Bezug auf Erfolgs- und Misserfolgsattribution, temperaments- und natürlich auch geschlechtsbedingte Unterschiede (siehe Scholz 2012, S. 10 ff.).

Die Mikro- und Makroebene von innerer Differenzierung in der Schule

Um der Heterogenität gerecht zu werden, erfolgt bei der Unterrichtsplanung die Anpassung des jeweiligen Unterrichts an die entsprechende Klasse, die sogenannte innere Differenzierung (Hänze 2009, S. 4). Hierbei sollen Lerngelegenheiten geschaffen werden, bei denen nicht alle Lernenden auf die selbe Weise, aber möglichst erfolgreich und befriedigend lernen können (Wodzinski und Wodzinski 2007, S. 6). Das kann zum Beispiel über das Arrangement von unterschiedlichen Zielen, Inhalten, Methoden, Medien oder Sozialformen geschehen (Stäudel 2009b, S. 10). Nach Ansicht des Ministeriums für Wissenschaft, Forschung und Kunst in Baden-Württemberg, sollte der gelungene Umgang mit Heterogenität dabei auf folgenden Ebenen in der Schule stattfinden:

- Auf der Makroebene sollen Maßnahmen auf Klassenebene sein. Diese beinhalten nicht die Aufteilung der Schüler entsprechend der äußeren Differenzierung auf Schulformen, sondern die schulinterne Gruppierung nach Leistungsniveaus. Ferner wird hier die Anpassung des Schulcurriculums und Zusatzangebote von kompensatorischen Trainings vorgeschlagen.
- Auf der Mikroebene sollen Anpassungen innerhalb des Unterrichts stattfinden. Das geschieht durch Steuerung der Lehrer-Schüler-Interaktionen, durch die Auswahl bestimmter Aufgaben und dem Angebot von zusätzlichen Erklärungen (Volkholz et al. 2013, S. 32).

Mögliche Abläufe von Differenzierung im Unterricht

Auf der Mikroebene, also der Ebene der konkreten Umsetzung von Unterricht, stellt Bohl fünf Abläufe von Unterricht vor. Diese Abläufe werden in Abb. 4 dargestellt, wobei sich der Buchstabe *L* auf die Lehrkraft und *A* auf die Aufgaben oder das Material bezieht.

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 4: Möglichkeiten von Unterrichtsabläufen auf der Mikroebene (entnommen aus Bohl 2012, Folie 18).

Zu den Abbildungen führt Bohl folgende Ergänzung auf (Bohl 2012, S. 20 Folie 20 ff.): Der individualisierte Unterricht erfasst die Lernvoraussetzungen einzelner Schüler und nimmt diese als Ausgangspunkt, um passende und individualisierte Lernangebote für den jeweiligen Schüler zu machen. Der individualisierte Unterricht ergibt aber nur dann Sinn, wenn die Voraussetzungen eines Schülers genau geklärt sind.

Durch diesen Prozess geht ein hohes Maß der Anpassung mit einem hohen Maß der Fremdbestimmung des Schülers einher.

Differenzierter Unterricht folgt dem gleichen Ansatz wie der individualisierte Unterricht, geht dabei allerdings nicht von einzelnen Schülern, sondern von den gleichen Bedürfnissen einer Schülergruppe aus. Im Gegensatz dazu geht offener Unterricht mit einem hohen Ausmaß von Selbst- und Mitbestimmungsmöglichkeiten der Schüler einher. Dies setzt allerdings eine gewisse Reife der Schüler und Übung im Umgang mit solch einer Form des Unterrichts voraus oder muss geübt werden. Dennoch treffen Schüler bei einem offenen Unterricht mehr eigenverantwortliche Entscheidungen in Bezug auf den organisatorischen, methodischen und inhaltlichen Ablauf ihres Unterrichts als bei den anderen Formen. Mit adaptivem Unterricht ist die Anpassung des Lernangebots an die individuellen Voraussetzungen der Lernenden gemeint. Die Form der adaptiven Instruktion kann sowohl ein individualisierender als auch ein differenzierender oder offener Unterricht sein. Die adaptive Instruktion ist demnach als eine Kombination dieser drei im Ablauf einer Unterrichtsstunde oder Einheit zu sehen, die den Voraussetzungen der Lernenden und der Planung der Lehrkraft entspricht. In Anbetracht dieser Unterrichtsformen zielt die vorliegende Forschungsarbeit auf die Unterstützung von Lehrkräften bei der Planung und Durchführung von differenzierendem Unterricht in Biologie ab.

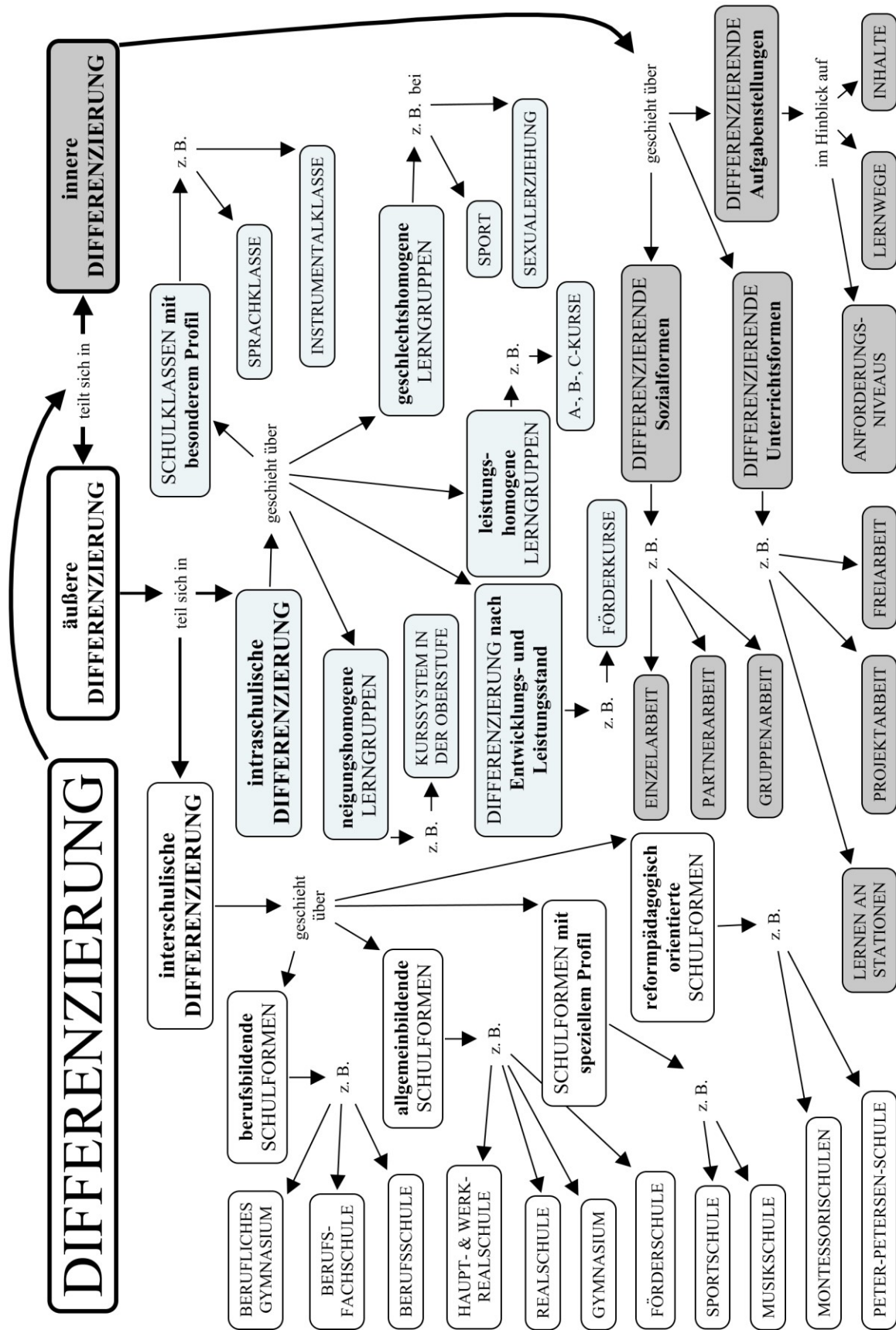


Abb. 5: Überblick über Differenzierungsmöglichkeiten in Deutschland (in Anlehnung an Scholz 2012).

3.2.2 Voraussetzungen für differenzierenden Unterricht

Merkmale und Kriterien für guten Unterricht sind hinlänglich bekannt und geben Hinweise zur Planung von lernwirksamem Unterricht (Meyer 2004, Helmke 2007). Die Vergleichsstudie von John Hattie (Hattie 2009) zeigt jedoch, dass manche Merkmale und Kriterien für sich alleine genommen keine großen Auswirkungen auf das Lernen haben. Aus diesem Grund kommt es auf das Zusammenwirken, der *Orchestrierung* von Unterrichtsmerkmalen, an (Pietsch 2010).

Besonders deutlich wird das Ausmaß des Zusammenspiels bei dem Merkmal des individualisierten Unterrichts. Für sich alleine und aus dem Zusammenhang gerissen hat diese Form der Unterrichts wenig Einfluss auf das Lernen ($d=0.23$, Hattie 2009). Wird im Unterricht also von heute auf Morgen zum Beispiel durch den Einsatz von unterschiedlichen Aufgaben individualisiert, ist die Chance auf Lernsteigerung gering, sofern die Änderung des Unterrichts nicht mit anderen lernwirksamen Merkmalen verknüpft wird (zum Beispiel klare Ziele, kognitive Aktivierung, adaptive Unterstützung). Damit differenzierender Unterricht gelingen kann, müssen bestimmte Kriterien im Unterricht erfüllt werden. Im Folgenden wird hierzu eine Studie vorgestellt, die anhand eines Stufenmodells die Voraussetzungen für differenzierenden Unterricht beschreibt.

Abstufung von Unterrichtsqualität

Basierend auf einem Modell zur Evaluation von Unterrichtsqualität durch Bildungsstandards nach Pietsch (Pietsch 2010) konnten vier Qualitätsstufen von Unterricht identifiziert werden. Für die Entwicklung des Qualitätsstufen-Modells wurden 30 Kriterien für die Beobachtung von gutem Unterricht aus aktueller schulpädagogischer und pädagogisch-psychologischer Literatur extrahiert (zum Beispiel Helmke 2003, Helmke 2006a, Meyer 2004). Diese Kriterien wurden anschließend mittels einer explorativen Faktorenanalyse auf sechs beobachtbare Kategorien aufgeteilt:

1. Klassenmanagement und Klassenklima,
2. Unterricht strukturieren, Methoden variieren,
3. motivieren, intelligent Üben, aktiv lernen,
4. Schülerorientierung und Unterstützung,
5. individuelle Förderung und
6. Lernerfolgssicherung.

Im Anschluss wurden 2240 Beobachtungen an unterschiedlichen Schulen und Schultypen durchgeführt. Die Ergebnisse wurden ausgewertet und in einem vierstufigen Modell zusammengefasst. Differenzierender Unterricht entspricht der vierten Stufe

und kann nur gelingen, wenn die drei vorhergehenden Stufen bereits durchlaufen wurden. Da die Stufen die Grundlage für das Gelingen von Binnendifferenzierung im Unterricht bilden, werden sie im Folgenden vorgestellt:

- Unterricht auf **Niveaustufe 1** zeichnet sich durch ein lernförderliches Unterrichtsklima aus. Die Unterrichtseinheiten folgen einer klaren Struktur, wobei die Lehrkraft bei Bedarf von der Planung abweicht und auf Schüler oder Situationen eingeht. Darüber hinaus werden vereinbarte Regeln eingehalten und Arbeitsaufträge klar, präzise und verständlich gestellt.
- Unterricht auf **Niveaustufe 2** zeichnet sich nicht nur durch die bisher genannten Merkmale aus, sondern ist auch gekennzeichnet von zielgerichteter Klassenführung, effektiv genutzter Lernzeit und einer Variation von zweckdienlichen Methoden im Unterrichtsgeschehen. Auf Stufe 2 werden bereits individuelle Leistungen der Schüler von der Lehrkraft wahrgenommen und gelobt.
- Auf **Niveaustufe 3** werden Schüler auf vielfältige Weise zum Lernen motiviert. Der Unterricht erfolgt schülerorientiert und weniger mechanisch. Der Unterricht zeichnet sich dadurch aus, dass Lernende zudem Freiheiten bekommen und mitbestimmen können. Auch wird ihnen ermöglicht, aktiv und selbstständig zu lernen, wobei ihre jeweiligen Lernstände anhand transparenter Verfahren gemessen und rückgemeldet werden.
- Abgesehen von Binnendifferenzierung zeichnet sich Unterricht auf **Niveaustufe 4** durch einen hohen Anspruch und den Einbezug von überfachlichen Zusammenhängen aus. Es kommt vermehrt zum Einsatz transparenter Diagnose- und Feedbackverfahren, wobei die Reflexion des eigenen Lernens und der eigenen Lernprozesse eine wichtige Rolle spielen. Schüler auf Stufe 4 des Unterrichts erhalten phasenweise auch die Möglichkeit, an selbstgewählten Lerninhalten zu arbeiten. Das führt zu einem nachhaltigen Kompetenzerwerb.

Möglichkeiten für differenzierenden Unterricht

Innere Differenzierung kann nach Pietsch also nicht einfach gemacht werden, sondern setzt einen guten Unterricht voraus. Sofern dies gegeben ist, könnten Lehrende ihren Schüler zum Beispiel schwierigkeitsbezogene Differenzierungsangebote machen, die verschiedene Bearbeitungswege zulassen oder durch Bearbeitungshilfen erleichtert werden (Bräu 2007, S. 178). Auch könnte nach Arbeitsweisen, Inhalten oder dem Lern- und Arbeitstempo differenziert werden (Leisen 2006, S. 265).

Differenzierungen können darüber hinaus durch Unterrichtsmaterialien erfolgen. Ausgewählte Möglichkeiten sind im Folgenden aufgeführt (entnommen aus Scholz 2008, S. 12 ff.):

- **Umfang des Lernstoffes:** Einige Schüler sind schneller als andere. Korreliert ihr Arbeitstempo nicht mit überdurchschnittlichen kognitiven Fähigkeiten, wollen sie in der Regel nicht anspruchsvollere, sondern eher mehr Aufgaben auf ähnlichem Niveau. Andernfalls langweilen sie sich, was nicht selten mit deutlicher Unruhe und entsprechenden Störungen des Unterrichts einhergeht. Um sie nicht für ihre Schnelligkeit zu bestrafen, sollte man ihnen motivierende Zusatzaufgaben mit spielerischem Charakter anbieten, wie zum Beispiel Rätsel, Lernspiele, Knobelaufgaben, interessante Zusatztexte, oder sie die entsprechenden Materialien selbst herstellen bzw. suchen lassen.
- **Anforderungsniveau:** Eine Differenzierung nach Leistungs- und Anforderungsniveau kann durch entsprechende Aufgaben oder Materialien erfolgen. Ergänzend oder alternativ zur leistungsdifferenzierten Aufgabenstellung kann ein differenziertes Materialienangebot zum Einsatz kommen. So bietet es sich gelegentlich an, zusätzlich zum Schulbuch zur fakultativen Vertiefung einzelner Aspekte weitere Materialien zur Verfügung zu stellen (Glossar, Lexika, weitere Texte u.a.), für die schwächeren Schüler zur Orientierung und Erleichterung den Schulbuchtext mit entsprechenden Hilfestellungen zu versehen und den leistungsstarken oder motivierten Schülern zusätzlich anspruchsvolle Originalquellen zuzumuten.
- **Inhalte und Interessen:** Manchmal können vorher im Plenum eingeführte Arbeitstechniken und Methoden, an verschiedenen Inhalten eingeübt oder vertieft werden. Dabei sollte man die Schüler individuell wählen lassen, an welchem Gegenstand sie die gestellte Aufgabe erarbeiten wollen. Die Differenzierung nach Interessen erhöht die Motivation und sorgt bei der abschließenden Präsentationsphase für entsprechende Vielfalt.
- **Lernstile und Zugangsweisen:** Für die Einschätzung der Lernstile eignen sich für die Unterrichts- und Schulpraxis entsprechende Fragebögen oder Tests, anhand derer Schüler ihre besonderen Stärken entdecken und gegebenenfalls nutzen können. Die Durchführung und Auswertung dieser Tests erfordert zum Teil nur einen geringen Zeitaufwand, so dass man sie ohne Probleme in einer Unterrichtsstunde bewältigen kann. Durch diese Tests erhalten die Schüler möglicherweise wertvolle Anregungen, auf welchem Weg sie sich Inhalte besonders gut aneignen können. In Tab. 1 auf der folgenden Seite sind die wichtigsten Lernwege mit Lösungsansätzen im Unterricht genannt.

Tab. 1: Differenzierungsmöglichkeiten über den Lernweg (entnommen aus Scholz 2008, S. 15).

Lernweg	Differenzierungsvorschlag
auditiv	z. B. Lehrervortrag, Audio-CD
haptisch	z. B. Vokabel- oder Textpuzzle, Legen einer geometrischen Figur
visuell	z. B. Abbildungen
handlungsorientiert	z. B. Gestaltung einer kleinen Spielszene
kognitiv-analytisch	z. B. Gestaltung einer Mind-Map
kommunikativ-kooperativ	z. B. Austausch mit einem Mitschüler über ein gestelltes Thema oder gegenseitiges Abfragen von Vokabeln

Aspekte für Differenzierungsmöglichkeiten als Orientierungshilfe

Angesichts der bereits dargestellten Masse an Möglichkeiten zur inneren Differenzierung ist ein Kategoriensystem zu deren Ordnung sinnvoll. Es erleichtert die Übersichtlichkeit und fördert die Verortung von einzelnen Differenzierungsmaßnahmen. Ein sehr einfaches als auch zweckdienliches System wird von Krüger vorgeschlagen. Hiernach sollten sich Lehrkräfte, die differenzierenden Biologieunterricht planen, folgende vier Fragen stellen: Wer soll differenzieren? Wonach soll differenziert werden? Was wird differenziert? Wie wird differenziert? Abb. 6 greift die vier Fragen auf und stellt gleichzeitig entsprechende Differenzierungsansätze dar (Krüger und Meyfarth 2009, S. 9).

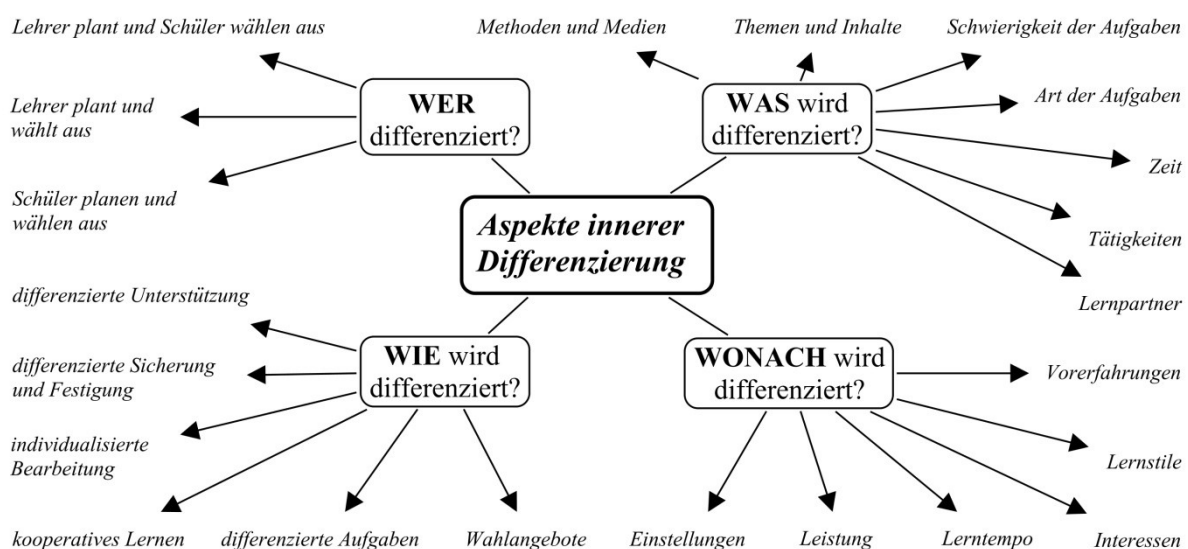


Abb. 6: Aspekte innerer Differenzierung in Anlehnung an Krüger (2009, Seite 9).

Im Vergleich zu den bereits erwähnten Lernmaterialien scheinen Lernaufgaben für die innere Differenzierung besonders gut geeignet zu sein (Wodzinski 2007, S. 16). So kann nach Bräu (2007) und Bohl (2012) überhaupt nur innere Differenzierung geschehen, wenn der Unterricht aufgabenorientiert ist. Entsprechend der Forschungsfragen und deren Wichtigkeit in Bezug auf die Materialentwicklung wird auf differenzierende Unterrichtsmaterialien wird im Kapitel 3.3 ab Seite 36 und auf Lernaufgaben in Kapitel 3.5 ab Seite 56 genauer eingegangen.

3.2.3 Kriterien von Differenzierung im Unterricht

Forschungsergebnisse zeigen (Köller et al. 2006), dass sich binnendifferenzierende Maßnahmen nicht nur auf Schülerleistungen auswirken, sondern dass auch Interesse und Selbstkonzept von Schülern eine Rolle beim Lernen spielen. So kann es durch Differenzierungsmaßnahmen innerhalb einer Klasse passieren, dass das Selbstkonzept bei leistungsstärkeren Schülern weiter steigt, wohingegen das Selbstkonzept von leistungsschwächeren Schülern sinkt (Schneider und Ludwig 2012, S. 72). Es kann auch gezeigt werden, dass gerade die Unterstützung von leistungsschwächeren Schülern oft auf Kosten leistungsstärkerer Schüler geht (Bohl 2012, S. 11). Leistungsschwächere Schüler lernen am besten in leistungsheterogenen Gruppen, wohingegen durchschnittliche Schüler vom Lernen in homogenen Gruppen profitieren (Tomlinson et al. 2003, S. 132). Dies ist ein klassischer Effekt von Aktion und Reaktion. Es hängt demnach sehr stark von der Planung und Durchführung differenzierender Maßnahmen ab, welche Auswirkungen diese haben (Schneider und Ludwig 2012, S. 72). Um den negativen Effekten vorzubeugen, werden im Folgenden Gütekriterien von innerer Differenzierung genannt.

Keine Differenzierung ohne Diagnose.

Wie bereits erwähnt entsprechen Schülerorientierung und produktiver Umgang mit Heterogenität den theoretischen und evidenzbasierten Qualitätsbereichen von gutem Unterricht (Helmke 2006b, Pietsch 2010). So sollte Unterricht aus konstruktivistischer Lernperspektive zum Beispiel theoretisch immer kognitiv aktivierend sein und einer gezielten Lehr-Führung entsprechen (Mayer 2004, S. 14). Denn wie bei einer video-basierten Analyse von Schülerarbeitsphasen gezeigt werden konnte (Krammer 2009, S. 7), ist gerade die adaptive Orientierung ein Kernmerkmal zur Unterstützung individueller Lern- und Verstehensprozesse. Von daher ist das wichtigste Qualitätskriterium die Anpassung der Differenzierungsmaßnahme an das Niveau des Schülers (Kalyuga et al. 2003, S. 29, Bohl 2012). Die Anpassung an das Niveau des Lernenden wiederum erfordert eine genaue Diagnose. Denn ohne fundierte Diagnose stoßen Differenzierungsmaßnahmen ins Leere, weil sie auf reiner Spekulation beruhen (Helmke 2006b,

S. 10). Etwas drastischer formuliert entsprechen unbegründete Differenzierungsmaßnahmen einer Therapie ohne Krankheitsbefund. Nach Helmke eignen sich für die Diagnose bereits bestehende standardisierte Instrumente, welche auf den Bildungsservern der Bundesländer zu finden sind. Als weitere Referenzen nennt er das Staatsinstitut für Schulqualität und Bildungsforschung (ISB) in Bayern, die Schweizer Arbeitsgemeinschaft Externe Evaluation von Schulen (ARGEV) und die österreichische Website SQA vom Bundesministerium für Unterricht).

Eine schulische Diagnose kann und soll aber keine klinische Diagnose sein. Aus pädagogischer Sicht sollte sich eine Diagnose nach Scholz (2012) nicht auf die Defizite von Schülern beschränken, sondern auch deren Stärken betonen. Ferner können im pädagogischen Kontext die psychologischen Gütekriterien (Objektivität, Reliabilität und Validität) nur bedingt gelten. Eine pädagogische Diagnose sollte sich Frau Scholzens Meinung nach an Weinert und Schrader (Weinert und Schrader 1986) orientieren und sich vor allem durch folgende Gütekriterien auszeichnen (Auflistung nach Scholz 2012, S. 24):

Eine pädagogische Diagnose sollte:

- eine ungefähre Genauigkeit aufweisen und permanent einer kritischen Reflexion unterzogen werden. Das besondere Augenmerk der Lehrer gilt dabei nicht nur dem aktuellen Leistungsstand, sondern vor allem der Frage, durch welche Faktoren der Lernprozess der Schüler beeinträchtigt und unterstützt werden kann.
- neben der sozial- und kriterienorientierten Bezugsnorm vor allem einen an den individuellen Fähigkeiten und Lernfortschritten orientierten Maßstab anlegen, der für den schulischen Lernerfolg und die Persönlichkeitsentwicklung der Schüler von großer Bedeutung ist.
- sich durch pädagogisch günstige Voreingenommenheit auszeichnen, das heißt Lehrkräfte sollen das Ausmaß der Leistungsunterschiede zwischen den Schülern einer Klasse maßvoll unterschätzen, das Leistungspotenzial einzelner Schüler leicht überschätzen und deren Erfolge auf Begabung und Misserfolge auf mangelnde Anstrengung oder ineffektiven Unterricht zurückführen. Diese leicht optimistische Erfolgserwartung stellt sicher, dass der Lehrer in seinen pädagogischen Bemühungen nicht nachlässt.

Schulische Diagnosen haben eine Schattenseite. Ausgehend von der Gefahr, dass sich Beurteilungen von Lehrern über die gleiche Schülerleistung gravierend unterscheiden können, sollten Schüler ihr Anspruchsniveau ab und zu auch eigenständig wählen dürfen. Dies soll der Stigmatisierung der Schüler entgegenwirken, denn eine gut gemeinte Hilfe durch eine Lehrkraft kann auf Dauer von Schülern etikettierend und abwertend

wahrgenommen werden: „einmal schwach, immer schwach“ (Scholz 2008, S. 13). In Anbetracht des sehr komplexen Ziels der Anpassung eines sowieso schon vielschichtigen Konstrukts namens Unterricht ist ersichtlich, dass es für die innere Differenzierung keinen allumfassenden Königsweg geben kann (Prediger 2008, S. 1). Wenngleich es bisher keine Gesetzmäßigkeiten für Differenzierung gibt, so formuliert Tomlinson generelle Prinzipien, nach welchen sich ein differenzierender Unterricht richten sollte (Tomlinson et al. 2003, S. 131 ff.). Hierfür schlägt sie folgende Leitlinien vor:

Ein guter differenzierender Unterricht:

- ist zielgerichtet,
- beinhaltet den flexiblen Einsatz von kleinen Lerngruppen im Klassenzimmer,
- variiert die Materialien, welche von Einzelpersonen und kleinen Gruppen im Unterricht benutzt werden,
- ist zeitlich flexibel und lässt unterschiedliche Bearbeitungszeiten zu,
- ist wissens- und lernerzentriert.

3.2.4 Veränderungen der Lehrerrolle durch Binnendifferenzierung

Während sich im Schulalltag auf der einen Seite allgemeine Bildungsstandards und zentrale Diagnose- und Vergleichsarbeiten etablieren, wird auf der anderen Seite die Forderung nach individueller Förderung der Schüler immer lauter (Neuhaus und Vogt 2005, S. 82). So müssen Lehrkräfte an allen Schulformen mit den individuellen Unterschieden innerhalb ihrer Schülergruppen adäquat umgehen können. Damit keine Risikogruppe zurückfällt, müssen Lehrer ihren Unterricht so gestalten, dass das Potential der Schüler optimal genutzt wird. Gleichzeitig müssen im Unterricht aber auch die Mindeststandards gesichert und gehalten werden (Volkholz et al. 2013, S. 32).

Neue Ansprüche durch Binnendifferenzierung gehen in der Praxis mit Problemen einher. Lehrer sollen ihre Schüler zwar fördern und individuelles Lernen ermöglichen, am Ende müssen sie ihre Schüler aber wieder durch Leistungssituationen beurteilen, was für Lehrer zu einer paradoxen Doppelrolle führt (Luthiger 2008, S. 39). Ferner gehen die Ansprüche zwangsläufig mit einer Veränderung des bisherigen Unterrichts einher. Die Vorstellung, den eigenen Unterricht jedoch einer externen Bestandsaufnahme (externe Evaluation) zu unterziehen und entsprechend zu verändern, ist für manche Lehrer ungewohnt und teilweise auch unangenehm, da manche Erneuerungen im Widerspruch zu langjährigen Gewohnheiten stehen können (Helmke 2006b, S. 11). Von daher ist mit Vorbehalten und Widerständen zu rechnen, die sich mit Ergebnissen der vorliegenden Studie decken (siehe Punkt 4.3.1 auf Seite 97).

Aber selbst dann, wenn Lehrer sich der nötigen Veränderung bewusst sind und sagen, dass sie etwas Bestimmtes in ihrem Unterricht umsetzen möchten, ist das noch keine Garantie für dessen tatsächliche Umsetzung. So konnte Polly (Polly und Hannafin 2011, S. 126) in einer Videoanalyse zeigen, dass (auch) bei Lehrern ein großer Unterschied zwischen Selbst- und Fremdwahrnehmung besteht. Dies zeigte sich auch bei Polly ganz besonders darin, wie Lehrer im Interview ihren Unterricht in Bezug auf Schülerzentrierung einschätzten und wie wenig das tatsächlich der Fall war (Polly und Hannafin 2011, S. 128).

Damit Veränderungen im Unterricht überhaupt ankommen, müssen Lehrer nach Aussage der Bund-Länderkommission (1997) persönliche Weiterentwicklung als berufliche Aufgabe akzeptieren. Diese kann durch Fortbildungen zwar unterstützt werden, allerdings spielt dabei die innerberufliche Kommunikation, der kollegiale Austausch und Kooperationsfähigkeit eine wesentliche, wenn nicht sogar die bedeutendste Rolle (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung 1997, S. 60). Die Pflege der eigenen Lehrerprofessionalität beinhaltet hierbei sowohl das fachliche Verständnis der Beziehungen der zu unterrichtenden Inhalte (Krammer 2009, S. 122), als auch eine wirksame Handlungskompetenz bei der Verknüpfung von Wissen und Können. Knapp 16 Jahre nach dem BLK-Gutachten wurde eine Expertenkommission zur Weiterentwicklung der Lehrerbildung in Baden-Württemberg befragt. In der Empfehlung der Expertenkommission wurden Kernkompetenzen professioneller Lehrkräfte in Bezug auf Wissen, Überzeugung, Motivation und berufliche Selbstregulation genannt (siehe Tab. 2).

Tab. 2: Aspekte professioneller Kompetenz von Lehrkräften (entnommen aus Volkholz et al. 2013, S. 33).

Aspekte professioneller Kompetenzen von Lehrkräften	Bereiche
Wissen	Fachwissen, fachdidaktisches Wissen, allgemeines pädagogisch-psychologisches Wissen, bildungswissenschaftliches Wissen, diagnostische Expertise als übergreifender Bereich, der fachspezifische und allgemeine Bereiche betrifft
Überzeugungen	Zum Selbst (eigene Rolle, Erziehungsziele), über Lernen und Lernprozesse, über den Bildungskontext (z. B. Strukturen/Reformen), über Bildung/Kindheit allgemein
Motivation	Selbstwirksamkeitsüberzeugungen, Ziele, Interessen, intrinsische/extrinsische Motivation
Berufliche Selbstregulation	Engagement, Umgang mit den eigenen Ressourcen, Selbstmanagement (z. B. Zeitmanagement), Kooperation, Teamfähigkeit, Selbstreflexion (Offenheit zur Rezeption neuer Aspekte und Informationen)

Besondere Anforderungen an Lehrkräfte durch Binnendifferenzierung

Ausgehend von den allgemeinen Kompetenzen von Lehrkräften und den Anforderungen, die aus einem differenzierenden Unterricht erwachsen, sollen in diesem Abschnitt die Kompetenzen identifiziert werden, die für Binnendifferenzierung besonders bedeutend sind. Wie bereits erwähnt gilt es, bei Maßnahmen der inneren Differenzierung den Unterricht von Schülern aus zu planen und durchzuführen. Hierbei wechselt die Rolle des Lehrers weg vom traditionellen Informationsvermittler hin zum Lernorganisator, welchem eher eine beratende als instruierende Funktion zukommt (Wodzinski und Wodzinski 2007, S. 9).

Nach Hellrung (2009) bedeutet das für Lehrkräfte, dass sie den Schülern einen Möglichkeitsraum für selbstreguliertes Lernen eröffnen sollen. Hieraus entwickeln sich besondere Anforderungen, die aus Binnendifferenzierung entstehen (siehe Tab. 3).

Tab. 3: Anforderungen an Lehrkräfte, die aus Binnendifferenzierung erwachsen (in Anlehnung an Hellrung 2009, S. 113).

Für einen differenzierenden Unterricht müssen Lehrkräfte...

1. den individuellen Lern- und Leistungsstand sowie individuelle Lernschwierigkeiten bzw. Hindernisse der Schüler erkennen,
2. den jeweiligen Unterstützungsbedarf der einzelnen Schüler erkennen und entsprechende Unterstützungsangebote machen,
3. sich in individuelle Lernwege eindenken,
4. individuell passende Lernstrategien vorschlagen,
5. den individuellen Lernprozess im Blick behalten und dokumentieren,
6. konstruktive Rückmeldungen zum Lernprozess und zur Leistungsentwicklung geben,
7. machbare Vereinbarungen zur Gestaltung des Lernprozesses mit jedem Schüler aushandeln.

Die Erfüllung dieser differenzierungsspezifischen Anforderungen muss sowohl auf der Makro- als auch auf der Mikroebene des Unterrichts geschehen. Sie darf nicht punktuell stattfinden, sondern muss kontinuierlich erfolgen. Sofern es dabei nötig wird, sollten die Erziehungsberechtigten und anderes Fachpersonal in den Prozess mit einbezogen werden, um Förderpläne abzustimmen, die über den Unterricht hinausgehen. In Abb. 7 werden zusätzliche Anforderungen an Lehrer dargestellt, die sich durch Binnendifferenzierung ergeben (Volkholz et al. 2013, S. 32 f.).

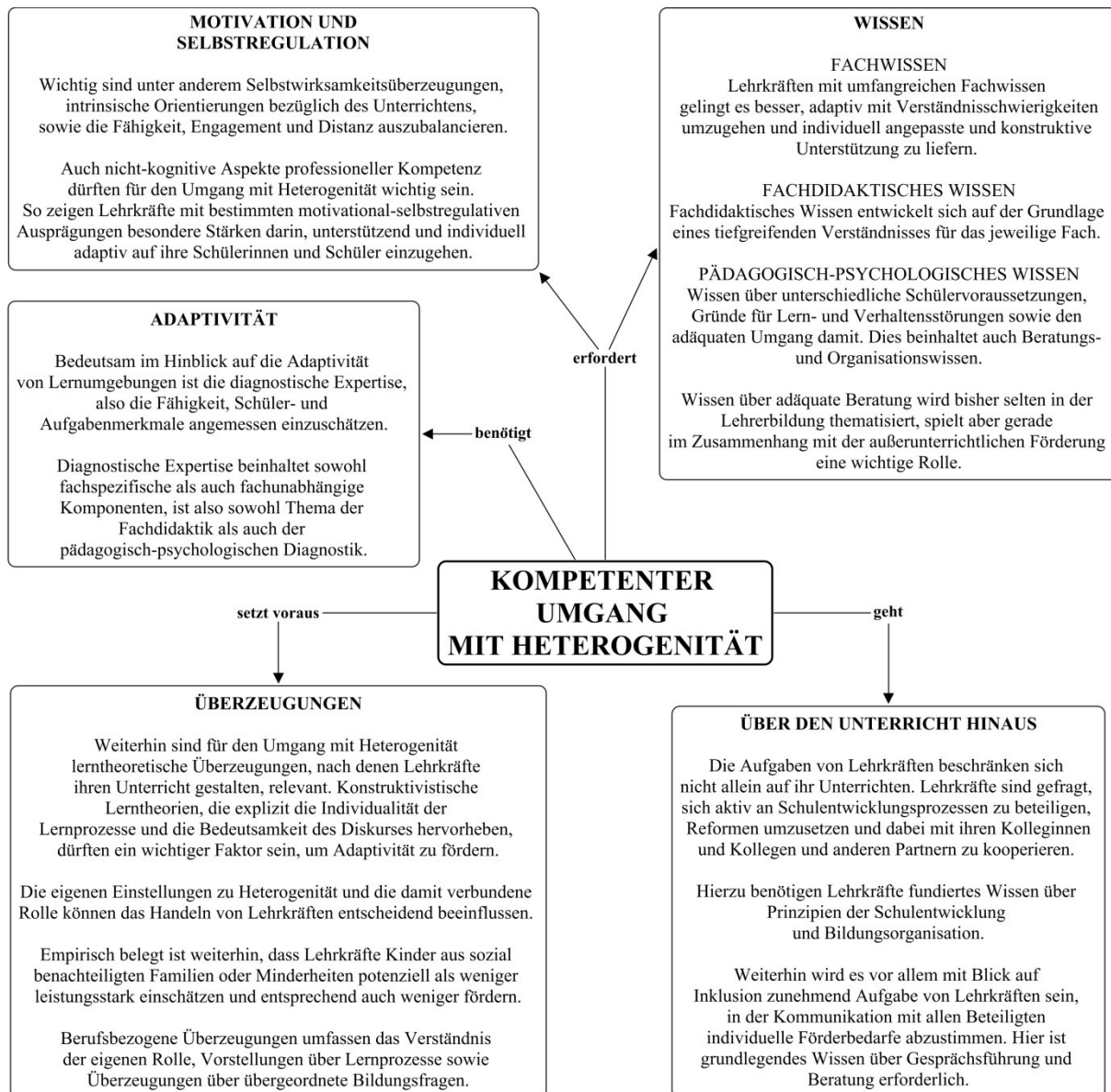


Abb. 7: Anforderungen an Lehrkräfte, die durch Binnendifferenzierung entstehen (in Anlehnung an Volkholz et al. 2013, S. 32 f.).

Nach der Aufführung der zusätzlichen Anforderungen werden folgende Punkte deutlich: Innere Differenzierung ist sehr komplex und geht mit hohen Anforderungen für Lehrer einher. In Bezug auf Differenzierung im Schulalltag zeigte sich dies schon in einer BLK-Studie (1997), welche erwies, dass Lehrer für die Planung von Unterricht meist bessere Kenntnisse in ihrem Fach hatten als für dessen methodische Umsetzung (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung 1997, S. 59).

Dieses Ergebnis wiederholt sich in aktuelleren Beiträgen zum Beispiel in Bezug auf die Konstruktion und den Einsatz von Aufgaben im Unterricht. Hierbei scheinen Lehrer anstatt begründet eher intuitiv wenn nicht sogar willkürlich vorzugehen (Winter 2008, S. 119, Luthiger 2008, S. 38). Auch die für die innere Differenzierung so wich-

tige diagnostische Kompetenz scheint im Alltag eher schwach ausgeprägt und vor allem Fehleranfällig zu sein. Scholz (Scholz 2012, S. 25) nennt hierfür Effekte, welche die Objektivität von Lehrkräften beeinflussen. In Tab. 4 werden diese Effekte aufgeführt und durch weitere ergänzt.

Tab. 4: Zusammenstellung subjektiver Fehlerquellen bei der Einschätzung von Schülern durch Lehrer. Hierbei handelt es sich um eine Zusammenstellung aus Sicht der Schulpraxis (teilweise nicht empirisch belegt, aber dennoch in der Praxis relevant).

Akteur-Beobachter-Effekt: Bei der Kausalerklärung einer Handlung einer Person unterscheidet sich die Begründung des Akteurs (beobachtete Person) von der des Beobachters systematisch. Beispiel: „Der an der Tafel geprüfte Schüler (beobachtete Person) kommt ins Stocken und schreibt dies den Umständen zu: Lehrer und Mitschüler machen ihn nervös, es ist die sechste Stunde. Dagegen weiß der Lehrer (beobachtende Person), dass dies nur Ausreden sind: Der hat einfach nicht gelernt!“³.

Beziehungs-Effekt: Sofern eine Lehrperson die Beziehung zu den Schülern über Sympathie definiert, läuft sie Gefahr, in Abhängigkeit zu geraten. Da eine schlechtere Note zu einer Verschlechterung der Beziehung führt, sorgt der Beziehungs-Effekt für mildere Noten³.

Erster Eindruck: Der erste Eindruck, den man von einem Menschen bekommt, ist oft sehr bestimmend für die weitere Wahrnehmung dieser Person. Er wird zum Bezugsrahmen für den weiteren Umgang und ist nur schwer zu korrigieren, obwohl er durch ganz spezifische und zufällige Umstände zustande gekommen sein kann. Beispiel: „Jeder der ein Bewerbungsgespräch vor sich hat weiß um die Bedeutung des ersten Eindrucks und bereitet sich darauf vor, einen möglichst guten Eindruck zu machen. Häufig entscheidet sich bereits in den ersten fünf Minuten, ob ein Bewerber eine Chance hat oder nicht“⁴.

Halo-Effekt: Auch Hof-Effekt genannt, tritt auf, wenn einzelne Eigenschaften den Gesamteindruck überstrahlen. Andersherum kann ein guter Gesamteindruck weniger gute Eigenschaften kompensieren⁵. Beispiel: „Wer sich gut ausdrücken kann, kann auch klar denken“ – Gute Erziehung und angepasstes Verhalten werden positiv belohnt und können die objektive Bewertung der Fachleistungen beeinflussen³.

Logischer Fehler: Bei diesem Beobachtungsfehler wird aus dem Vorliegen eines Merkmals auf das Vorliegen eines zweiten Merkmals geschlossen. Bei diesem Prozess werden Eigenschaften als logisch zusammengehörig empfunden. Beispiel: Ein ordentlicher Schüler ist auch fleißig, obwohl dafür keine Hinweise vorliegen. Im Gegensatz zu der diffusen Übertragung eines Merkmals auf andere (Halo-Effekt), werden beim logischen Fehler assoziative Merkmale übertragen, welche nicht oder nur bedingt vorliegen⁵.

³ Studienseminar Koblenz 2011

⁴ Prändl 2011

⁵ Dorsch et al. 2014

Projektion: Bei der Projektion werden emotionale Vorgänge innerhalb des Lehrers nach außen verlagert. Hierbei können äußere Wahrnehmungen durch innere Auseinandersetzungen überlagert werden. Dies kann sich in der Empfindung oder dem Erleben äußern, indem die subjektive Qualität des eigenen Gemütszustandes als Eigenschaft äußerer Dinge betrachtet oder auf die Einschätzung anderer Personen übertragen wird⁵.

Pygmalion-Effekt: Dieser Effekt beschreibt, dass die Erwartung einer Lehrkraft gegenüber einem Schüler tendenziell dessen Handlungen bedingen. Das bedeutet, dass die vom Lehrer gehegte Erwartung (Einstellung, Überzeugung und Vorurteil) gegenüber einem bestimmten Schüler das vom Schüler gezeigte Verhalten bzw. die Wahrnehmung des Schülerverhaltens durch den Lehrer tatsächlich beeinflusst. Dies kann dazu führen, dass die Leistungen eines für intelligent gehaltenen Schülers eher als gut, die Leistungen eines für nicht intelligent gehaltenen Schülers eher als schlecht eingeschätzt, bzw. wahrgenommen wird⁵.

Reihenfolge-Effekt: Bei der Bewertung von Klassenarbeiten spielt die Reihenfolge eine Rolle. So kann es alleine durch die Reihenfolge der Korrektur bei der ersten im Vergleich zur letzten Korrektur zu unterschiedlichen Bewertungen kommen, welche nicht auf den objektiven Leistungen der Schüler beruhen³.

Sympathie-Effekt: Lehrer tendieren dazu, für sie sympathisch wirkende Schüler unabhängig von deren Leistung besser zu bewerten, als Schüler, die ihnen unsympathisch sind⁶.

Zusatzinformationen: Sowohl eine positive als auch eine negative Information, die eine Lehrkraft über einzelne Schüler erfährt, kann Auswirkung auf dessen Benotung haben³. Im negativen Sinn bedeutet dies eine Bevorteilung oder Benachteiligung eines Schülers. Im positiven Sinne entspricht dies der pädagogischen Verantwortung von Lehrkräften. „Lehrkräfte sind bei der Vergabe von Zeugnisnoten nicht strikt an die sich aus den mündlichen und schriftlichen Leistungen ergebende rechnerische Gesamtnote gebunden. Sie haben bei der Notenvergabe in pädagogischer Verantwortung eine Gesamtbewertung vorzunehmen, welche die Beobachtungen im Unterricht sowie die Lern- und Leistungsentwicklung berücksichtigt“⁷.

⁶ Paradies et al. 2007

⁷ VG Braunschweig, Urteil vom 01.09.2010, S. 1

Lehrer brauchen Unterstützung bei Binnendifferenzierung

Um es zusammenfassend zu sagen: Lehrer brauchen für qualitativ hochwertigen differenzierenden Unterricht Fach- und Methodenwissen, fundierte lernpsychologische Kenntnisse und diagnostische Kompetenzen (Polly und Hannafin 2011, S. 129). In Bezug auf Bildungsreform, Binnendifferenzierung und all den damit verbundenen Anforderungen ist die Planung von entsprechendem Unterricht von einzelnen Lehrpersonen fast nicht mehr zu leisten (Bresler 2007, S. 33). Auch die Unterstützung von Lehrern durch entsprechend aufbereitete Materialien und Handlungsanweisungen, die dann nur noch eingesetzt und durchgeführt werden müssen, ist nur bedingt sinnvoll. So können Lehrern nicht einfach Bildungsstandards mit Aufgabenbeispielen und entsprechenden Diagnoseinstrumenten geliefert werden. Diese Materialien und zugrundeliegenden Prinzipien müssen zuerst verstanden werden, bevor sie gewinnbringend umgesetzt werden können (Neuweg 2008, S. 89). Die Veränderung von tief sitzenden Überzeugungen und Handlungsrouninen bei Lehrern benötigt allerdings Zeit und erfordert lang anhaltende Lern- und Reflexionsprozesse (Gräsel et al. 2004, S. 133). Dies kann ausschließlich durch Fortbildungen geschehen, die unterrichtsnahe, fachdidaktisch gestützte Inhalte thematisieren und einen eindeutigen Praxisbezug haben (Reyer et al. 2004, S. 196).

Die logische Konsequenz daraus ist die stetige Verzahnung zwischen Theorie und Praxis. Diese Zusammenarbeit auf gleicher Augenhöhe zwischen Lehrern und Bildungsforschern ist dabei für beide Seiten profitabel. Zum einen existieren im deutschsprachigen Raum zum Beispiel kaum Untersuchungen, welche die Einstellungen von Biologielehrern zum Fach und zum Unterricht untersuchen (Neuhaus und Vogt 2005, S. 76). Zum anderen ist die Forschung für die Implementation ihrer Ergebnisse auf die aktive Unterstützung und Weiterentwicklung durch Lehrkräfte angewiesen (Gräsel et al. 2004, S. 149). Für Lehrer bedeutet die Kooperation hingegen eine persönliche Erweiterung und Verbesserung ihrer Professionalität und damit ihres Unterrichts (Klieme und Rakoczy 2008, S. 224). Nur aus der Kombination dieser drei Elemente (Praxis, Theorie und Empirie) kann sich eine effektive Weiterentwicklung von Unterricht ergeben (Müller und Helmke 2008, S. 33).

Diese Forderung lässt sich bestätigen. Basierend auf der Hattie-Studie (Hattie 2009) wurde die Wirksamkeit von Interventionen im schulischen Kontext untersucht und gegenübergestellt. Hierbei hat sich gezeigt, dass sich fachspezifische Lehrerfortbildungen nachhaltiger auf das schulische Lernen auswirken als andere Faktoren. Die Ergebnisse sind in Tab. 5 dargestellt.

Tab. 5: Die Bedeutung von unterschiedlichen Faktoren für erfolgreiches schulisches Lernen: starke und sehr starke positive Effekte (in Anlehnung an Hattie 2009 entnommen aus Köller 2012, S. 3).

Das hilft bei der Verbesserung von Unterricht...	Zusatzangebote für starke Schüler $d = .39$
	Regelmäßige Leistungsüberprüfung (Tests) $d = .46$
	Direkte Instruktion $d = .59$
	Time on Task (genutzte Unterrichtszeit) $d = .59$
Das hilft noch besser...	Problemlösender Unterricht $d = .61$
	Fachspezifische Lehrerfortbildung $d = .67$
	Programme zur Leseförderung $d = .67$
	Lehrerfeedback $d = .72$

Da fachspezifische Lehrerfortbildungen ein nachgewiesener Faktor für erfolgreiches Lernen an der Schule sind (siehe Tab. 5), hatte das vorliegende Projekt von Anfang an den Anspruch, Lehrkräfte aktiv in den Entwicklungsprozess einzubinden und an der Forschungsarbeit zu beteiligen. Die Beschreibung der Teilnehmenden an der Fortbildung, sowie deren Ablauf und Ergebnisse werden in Kapitel 4 ab Seite 89 dargestellt.

3.2.5 Zusammenfassung

Differenzierung ist nicht nur nötig, sondern auch Pflicht (siehe Seite 12). Den Unterricht den Anforderungen der Schüler anzupassen entspricht nicht nur den Prinzipien von gutem Unterricht sondern ist auch ein zentraler Bestandteil zur Verbesserung von Schülerleistungen. Dabei sollte nicht nur allgemein, sondern auch fachlich differenziert werden. Bevor jedoch produktiv im Unterricht differenziert werden kann, müssen bestimmte Voraussetzungen erfüllt sein.

Die erste Voraussetzung lautet, dass im Unterricht nicht von heute auf morgen differenziert werden kann. Vielmehr muss der bestehende Unterricht auf bereits etablierten Strukturen beruhen und auf gutem Unterricht aufbauen. Nach dem Unterrichtsmodell von Pietsch (2010), das innere Differenzierung explizit beinhaltet, besteht Unterricht aus vier Stufen, wobei die höchste Stufe unter anderem Binnendifferenzierung enthält (siehe 3.2.2 ab Seite 19).

Die zweite Voraussetzung für Binnendifferenzierung ist ein begründeter Bedarf, der auf einer objektiven Diagnose beruht. Diese kann im schulischen Kontext nicht den Kriterien psychometrischer Testverfahren entsprechen, sollte allerdings so objektiv wie möglich sein. Hierfür müssen sich Lehrkräfte subjektiven Fehlerquellen bei der Einschätzung von Schülerleistungen bewusst sein, um ihre Bewertungen immer wieder reflektieren zu können (siehe Tab. 4 auf Seite 30).

Die beiden bisher genannten Voraussetzungen gehen mit speziellen Anforderungen an Lehrkräfte einher (siehe Abb. 7 auf Seite 29). Auch hier ist eine sofortige Umstellung auf differenzierenden Unterricht nicht möglich, denn die veränderte Lehrerrolle muss zuerst verstanden und eingeübt werden, bevor sie überhaupt gewinnbringend im Unterricht umgesetzt werden kann. Dies bedarf Zeit und persönliche Weiterbildung sowie externer Unterstützung durch Fortbildungen, Materialien und Methoden.

Die Steigerung von Binnendifferenzierung im Unterricht ist eine Umstellung, die weit über die Verwendung von Handreichungen oder den hier entwickelten Materialien hinausgeht. Binnendifferenzierung erfordert die Überprüfung des (eigenen) Unterrichts in Bezug auf die Niveaustufen und deren qualitativer Ausprägung. Sofern sie den Anforderungen für gelungene Differenzierung im Unterricht entsprechen und es einen begründeten Anlass für Differenzierungsmaßnahmen gibt, können die hier entwickelten Materialien im alltäglichen Unterricht zum Einsatz kommen. Im Vergleich zu den vielen Möglichkeiten von Differenzierung im Unterricht (siehe Abb. 6 auf Seite 22), entsprechen die in dieser Arbeit entwickelten Materialien den theoretisch fundierten Prinzipien (siehe Abb. 8 auf Seite 35). Innere Differenzierung bei der Aufgabenkonstruktion erfolgt hierbei über das Leistungsniveau der Schüler, indem unterschiedlich schwere Aufgaben, Hilfen und unterschiedliche Medien unter Berücksichtigung einer flexiblen Bearbeitungszeit gegeben werden.

3.3 Lehren und Lernen mit Repräsentationen

In diesem Kapitel geht es um die Entwicklung und den Einsatz von Medien im Unterricht. Zu Beginn wird eine Einführung über Medien aus Sicht der Lernpsychologie gegeben. Diese basiert auf Lehr- und Standardwerken der Psychologie. Im Anschluss daran wird die innere Differenzierung aus lerntheoretischer Sicht begründet.

Von Pädagogen und Didaktikern werden im Unterricht eingesetzte Materialien als Medien bezeichnet, Psychologen nennen dies *externe Repräsentationen*. Bei Repräsentationen handelt es sich um ein Konstrukt der Psychologie zur Beschreibung von mentalen Prozessen. Repräsentationen können grob in interne und externe Repräsentationen unterschieden werden. Das hängt davon ab, ob sie internalisiert (internal) sind oder von außen (external) an eine Person herangetragen werden.

Vereinfacht dargestellt beschreibt die interne Repräsentation eine bestehende Vorstellung eines Objektes oder einer Funktion im Gehirn einer Person. Diese interne Repräsentation kann sowohl beschreibend als auch bildlich vorliegen. Die gleiche verbale (descriptive) oder bildliche (depictive) Kodierung liegt auch bei externen Repräsentationen vor. Hierbei handelt es sich um wahrnehmbare Zeichen, die ein bestimmtes Objekt oder einen Prozess symbolisieren (Schnotz und Bannert 2003, S. 143, Weidenmann 2006). Je weiter (abstrakter) eine Repräsentation dabei von der Realität entfernt ist, desto besser muss dem Schüler die Kodierung bekannt sein, um von der externen Repräsentation eine Vorstellung zu internalisieren. Ein Überblick unterrichtsrelevanter Ebenen externer Repräsentationen in Abhängigkeit zu deren Abstraktionsniveau ist in Abb. 9 dargestellt.

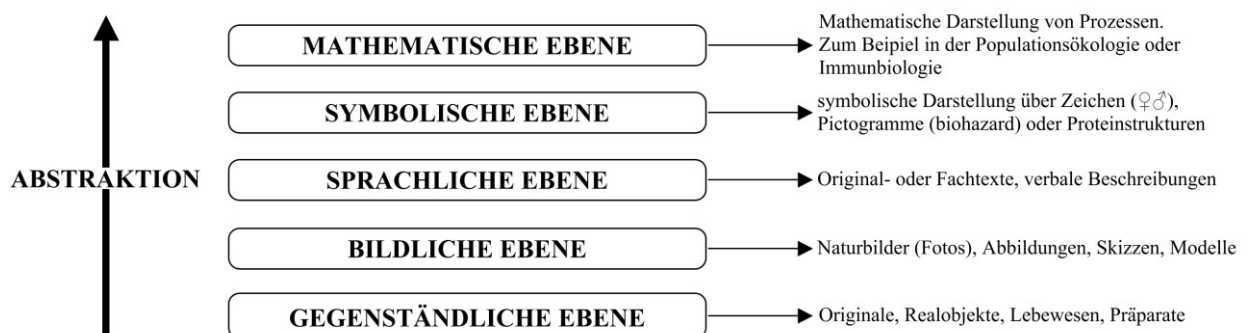


Abb. 9: Ebenen externer Repräsentationen im Biologieunterricht in Anlehnung an Leisen (2005, S.8).

Zur Entwicklung von lernwirksamen Materialien und Methoden im Umgang mit Heterogenität im Biologieunterricht, werden, in den folgenden Abschnitten lerntheoretisch relevante und anerkannte Ansätze zur Verarbeitung von Informationen durch unterschiedliche Repräsentationsformen dargestellt. Ausgehend von einer allgemein instruktionspsychologischen Vorstellung (Cognitive Load Theory) wird im Anschluss die Wirkung von Repräsentationen auf das Lernen beleuchtet. Die aus den Lerntheorien ableitbaren allgemeinen Merkmale für lernwirksame Medien werden durch weitere Kriterien zur Erstellung von Lehrmaterialien von Horz und Schnotz ergänzt (2010). Horz und Schnotz werden hier aufgeführt, da sie über die allgemein lernförderlichen Eigenschaften von Lehrmaterialien hinaus den Schüler als Individuum in der Planung von Lerngelegenheiten berücksichtigen und entsprechende Hinweise auf die Anpassung von Lehrmaterialien in Bezug auf innere Differenzierung geben.

In Anbetracht der Vielzahl an externen Repräsentationsformen im Biologieunterricht wird der Fokus bei der Ableitung von Gütekriterien zur Gestaltung von Lehr-Lern-Medien ausschließlich auf Texte, Concept-Maps und Abbildungen gelegt.

3.3.1 Cognitive Load Theory nach Chandler und Sweller

Die Cognitive Load Theory von Sweller und Chandler (1991) geht davon aus, dass unser Gehirn beim Lernen belastet wird. Im Vergleich zum Langzeitgedächtnis ist die Kapazität des Arbeitsgedächtnisses jedoch begrenzt und stellt den limitierenden Faktor beim Lernen dar. Chandler und Sweller unterteilen diese Belastung (cognitive load) des Arbeitsgedächtnisses nochmals in drei Kategorien, welche sich explizit auf die Eigenschaften von Lehr-Lern-Materialien beziehen lassen. Die Kategorien werden in Abb. 10 dargestellt und im darauffolgenden Text erklärt.

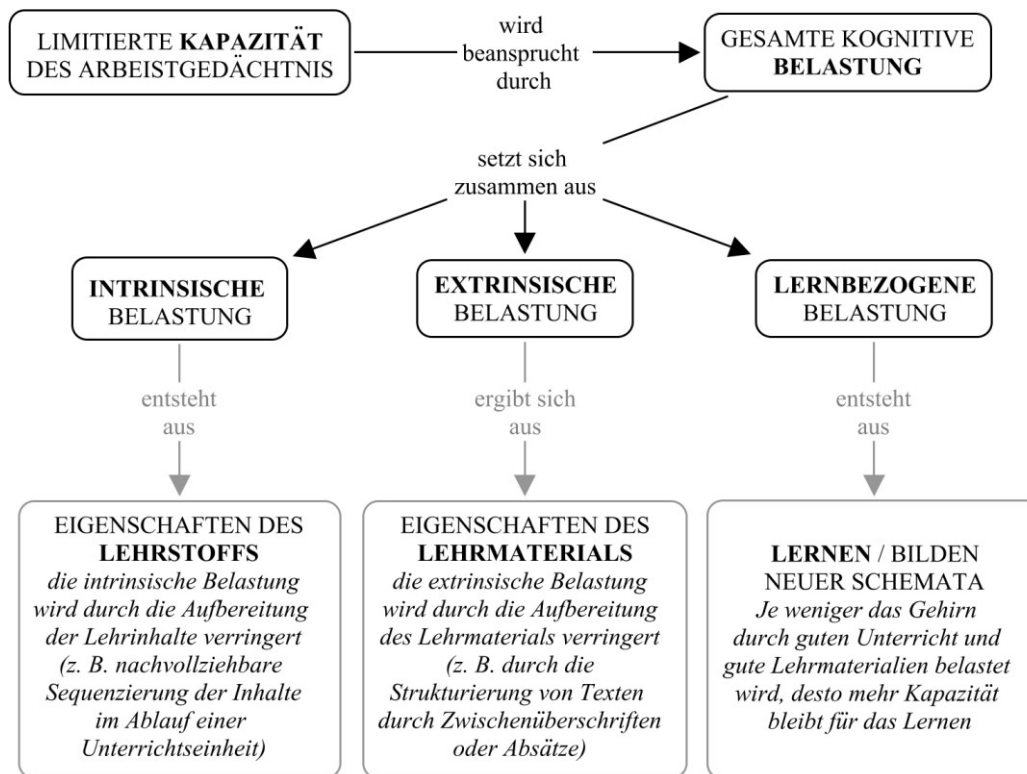


Abb. 10: Darstellung kognitiver Belastungen durch Unterricht und Lehrmaterialien (in Anlehnung an Sweller et al. 1998).

Die intrinsische kognitive Belastung (intrinsic cognitive load) entspricht der Belastung, die durch den Lernstoff selbst entsteht. Hierbei spielt die Abhängigkeit der Lerngegenstände untereinander oder wie viele davon zur gleichen Zeit bedacht werden müssen eine große Rolle. Als Beispiel sei hier das Lernen von Tieren im Wald und deren Eigenschaften genannt. Da Lernen isoliert und sequenziert geschehen kann (ein Tier nach dem anderen), ist die Belastung gering und kann mit Pauken verglichen werden. Versucht man die neu erlernten Tiere jedoch aus ökologischer Sicht zu betrachten und untereinander in Beziehung zu setzen, so wird die Sache komplexer und anspruchsvoller. Welche Konsequenzen hat zum Beispiel die Ausrottung einer Tierart für alle anderen Tiere? Was passiert, wenn es von der einen Tierart plötzlich weniger gibt oder wenn die Bäume sterben? Zusammenfassend gilt: Je komplexer der Lernstoff oder die Verbindungen der Lernelemente untereinander, desto höher die intrinsische Belastung (Sweller et al. 1998, S. 260). Die Einschätzung der intrinsischen Belastung durch den Lehrinhalt lässt sich allerdings nicht auf alle Personen in gleicher Weise übertragen. Denn Menschen mit bereits bestehendem Wissen auf einem Themengebiet, Sweller nennt es Schemata, fällt das Dazulernen leichter als Menschen ohne Vorwissen. Das Vorwissen ist dementsprechend ein kritischer Faktor für das Lernen. Je weniger Vorwissen vorhanden ist, desto schwieriger ist das Lernen.

Die extrinsische kognitive Belastung (*extraneous cognitive load*) entsteht aus den Eigenschaften des Lernmaterials an sich. Der Zugang zu relevantem Wissen kann alleine über die Gestaltung oder die Darstellung des Lernmaterials gesteuert werden. Hierzu gehört zum Beispiel überflüssige Zeilen aus Texten zu entfernen und die Texte auf Verständlichkeit, innere Kohärenz und Sequenzierung zu überprüfen (Krapp 2006, S. 444). Auf diese Weise sinkt der Aufwand bei der Bearbeitung des Textes. Der Blick auf das Wesentliche wird erleichtert und die extrinsische Belastung durch das Material reduziert.

Die lernbezogene kognitive Belastung (*germane cognitive load*) entspricht der tatsächlichen Belastung, die nur durch das Lernen entsteht. Sie beschreibt den kognitiven Aufwand einer Person, neues Wissen in einem Themengebiet aufzubauen oder mit bestehendem Wissen zu verknüpfen. Im Gegensatz zu der intrinsischen und der extrinsischen Belastung sollte die lernbezogene kognitive Belastung hoch sein. Wird das Arbeitsgedächtnis nämlich als begrenzt angesehen, sollte so wenig wie möglich von dieser begrenzten Kapazität für die Entschlüsselung von Lehrmaterialien oder die Entwirrung von Sachinhalten verwendet werden. Fazit: Je weniger das Arbeitsgedächtnis intrinsisch und extrinsisch belastet wird, desto mehr Kapazität kann für das Lernen genutzt werden (Sweller et al. 1998, S. 264). Für die Schule bedeutet das, dass der Lehrstoff immer entsprechend des Vorwissens der Schüler aufbereitet werden muss, indem er zum Beispiel in kleine Einheiten aufgeteilt und in eine logische, in sich schlüssige und aufeinander aufbauende Reihenfolge gebracht werden muss. Ferner bedeute dies, dass die Lehrmaterialien so geschaffen sein müssen, dass sie den Zugang zum Inhalt erleichtern und nicht erschweren. Wie das geht, wird in den folgenden Abschnitten geklärt.

3.3.2 Theorie zum Lernen mit Medien nach Mayer

Ähnlich der *Cognitive Load Theory* versucht Mayer bei seiner *Cognitive Theory of Multimedia Learning* die Verarbeitung von Information beim Menschen zu erklären. Hierbei wird im Vergleich zu Horz und Schnotz (2010) mehr Aufmerksamkeit auf das Lernen mit unterschiedlichen Medien (*multi media*) und kognitive Wahrnehmung gelegt. Bei dieser Arbeit geht es um den Vergleich zwischen einem Text und einer Experten-Concept-Map als Input. Da sich diese ausschließlich in ihrer Form unterscheiden, ist die *Cognitive Load Theory* als Lerntheorie unzureichend. Weil es in diesem Fall nicht um die Struktur der Inhalte sondern nur um den direkten Vergleich von zwei Repräsentationen auf das Lernen bei Schülern geht, beschreibt die kognitive Lerntheorie mit zwei medialen Inhalten den Lernprozess adäquater.

Die kognitive Lerntheorie mit multimedialen Inhalten nach Mayer (Mayer 2001, S. 44) geht von drei Grundannahmen aus:

- 1) Die Verarbeitung von Information geschieht über mehrere Verarbeitungskanäle. Diese Annahme beruht auf Paivios Dual-Code-Theorie, die besagt, dass Menschen auf (mindestens) zwei unterschiedlichen Wegen, nämlich Hören und Sehen, Informationen aufnehmen (Paivio 1986).
- 2) Die Verarbeitung von Information hat eine begrenzte Kapazität. Dies bedeutet, dass Eindrücke aus der Umwelt zwar über mehrere „Kanäle“ aufgenommen werden können, für deren gleichzeitige Verarbeitung jedoch nur limitierte mentale Ressourcen zur Verfügung stehen (Chandler und Sweller 1991).
- 3) Die Verarbeitung von Information ist kein passiver, sondern ein aktiver Prozess. Dabei wählen Personen gezielt relevante Informationen aus, organisieren sie und packen sie in eine schlüssige Vorstellung, die sie mit Vorerfahrungen und bereits Gelerntem verknüpfen (Mayer 2001 und Wittrock 1989).

Basierend auf diesen drei Annahmen entwirft Mayer (Mayer 2001, S. 42 ff.) sein Modell zum multimedialen Lernen. Er unterscheidet dabei drei Gedächtnisformen: das sensorische Gedächtnis, das Arbeitsgedächtnis und das Langzeitgedächtnis. Je nach medialem Repräsentationsformat gelangen äußere Eindrücke über die jeweiligen Sinnesorgane in das sensorische Gedächtnis. Hier erfolgt eine Selektion der Eindrücke, wobei nur eine Auswahl von Reizen in das Arbeitsgedächtnis gelangen. Diese Reize werden organisiert und entsprechend ihrer Art in eine verbale oder bildhafte Vorstellung umgewandelt. Diese Selektion⁸, Organisation, Reduktion und Umwandlung aller Eindrücke durch die Sinneskanäle erfolgt aufgrund der limitierten Kapazität des Arbeitsgedächtnisses und ermöglicht erst den wichtigsten Schritt des Lernens: die Integration. Hierbei werden zusammenhängende Elemente ungeachtet ihrer Art (verbal oder pictoral) miteinander verbunden und in bereits bestehenden Vorstellungen eingebunden. Von Lernenden erfordert dies Übung und anhaltende Aufmerksamkeit, bis aus einem Sinneseindruck eine im Langzeitgedächtnis gespeicherte, verknüpfte und bei Bedarf abrufbare Information geworden ist. Aus diesem Grund fordert Mayer in seiner Theorie des multimedialen Lernens die Erleichterung der Integration durch die Befolgung einfacher Prinzipien bei der Gestaltung von Repräsentationen.

⁸ Der Vorgang der Selektion eines Wortes bezieht sich immer auf einen Aspekt von dessen Bedeutung und nicht auf das Wort an sich. Genauso werden Bilder nicht komplett sondern nur Teilaspekte von diesen ausgewählt (Mayer und Moreno 2003).

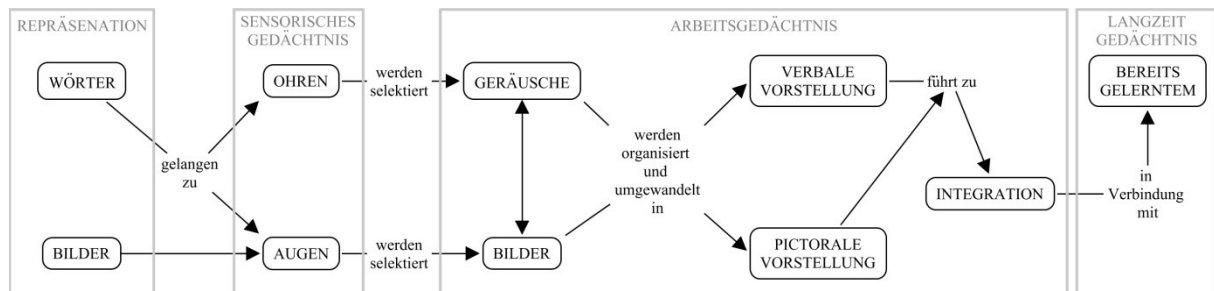


Abb. 11: Kognitive Lerntheorie nach Mayer (2001).

Nach Mayer sollten wenn möglich Text und Bild immer miteinander kombiniert werden, da diese beiden Repräsentationen im Verbund bei Lernenden mehr erreichen als getrennt. Ferner sollten sich gegenseitig ergänzende Texte und Bilder räumlich nahe beieinander platziert oder zumindest in zeitlicher Nähe nacheinander dargestellt werden. Gerade für leistungsschwächere Schüler empfiehlt Mayer, sich besonders Mühe bei der Gestaltung von Texten und Bildern zu geben, da diese im Vergleich zu leistungstärkeren Schülern stark davon profitieren. Trotz all der Mühe bei der Gestaltung von Medien rät Mayer jedoch dringend vom Einsatz schmückender und unterhaltender Bilder ab, da sie ablenken und das Arbeitsgedächtnis unnötig belasten.

3.3.3 Gütekriterien für Lehrmaterialien nach Horz und Schnotz

Basierend auf der Cognitive Load Theory und dem multimedialen Lernen von Mayer, gehen Horz und Schnotz (Horz und Schnotz 2010, S. 242) einen Schritt weiter. Der generelle Nutzen und die allgemeinen Eigenschaften von Repräsentationsformen sind bewiesen und bringen im sinnhaften Verbund mehr als einzeln (Mayer 2001). In den bisher aufgeführten Theorien liegt die Betonung auf dem bewussten Einsatz und der Erstellung von Repräsentationsformen, allerdings wird nicht genau gesagt, wie dies geschehen soll. Darüber hinaus hat sich gezeigt, dass generelle Aspekte der Cognitive Load Theory und die Theorie des multimedialen Lernens durchaus praxistauglich sind, jedoch nicht in allen Fällen zutreffen. So kann sich die Kombination von Text mit Bild aufgrund des Expertise Reversal Effect für Schüler mit Vorwissen negativ auf das Lernen auswirken (Kalyuga et al. 2003) und Schüler mit begrenzten räumlichen Fähigkeiten sogar überfordern (Plass et al. 2003). Nichtsdestotrotz sollte nach Horz und Schnotz die sinnvolle Kombination von Repräsentationsformaten das erste Prinzip für die Entwicklung von Lehrmedien sein (Horz und Schnotz 2010).

Als zweites Prinzip wird die Verarbeitungstiefe aufgeführt, die auf Mayer beruht (Horz und Schnotz 2010, S. 238). Hierbei entspricht eine geringe Verarbeitungstiefe (Lower Order Processing) dem Transfer von einem Reiz über ein Sinnesorgan in eine verbale oder pictorale Vorstellung im Arbeitsgedächtnis. Dieser Vorgang läuft automatisch ab und ist geringfügig bewusst zu beeinflussen. Im Gegensatz dazu stehen Vorgänge mit

einer hohen Verarbeitungstiefe (Higher Order Processing). Hierzu gehört sowohl die semantische (bedeutungsgemäße) Auswahl und Verarbeitung von Informationen und die Bildung von Zusammenhängen des betreffenden Lerngegenstandes als auch die Verknüpfung von neuen Informationen mit thematisch verwandten Inhalten im Langzeitgedächtnis (siehe Abb. 12).

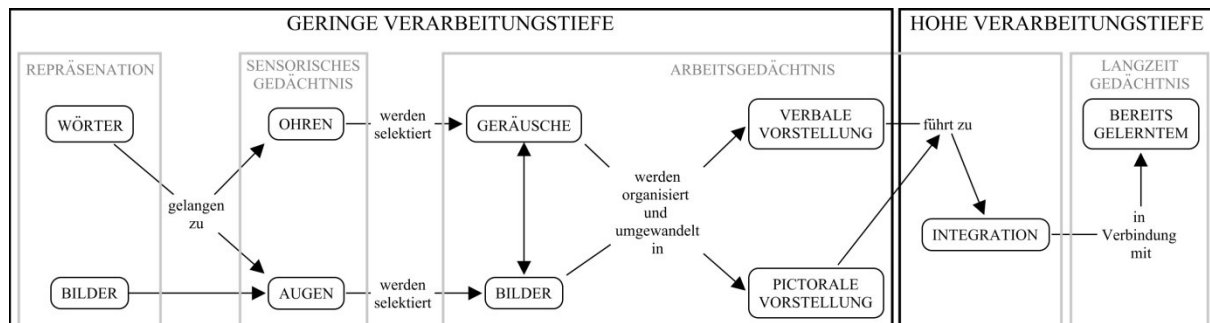


Abb. 12: Verarbeitungstiefe basierend auf Horz und Schnotz (2010).

Um die Entnahme von Informationen aus Lehrmaterial zu erleichtern und die verbale beziehungsweise picturale Vorstellung zu vereinfachen, sollten Kombinationen von Repräsentationsformen sowohl räumlich (spatial) als auch zeitlich (temporal) nahe beieinander liegen. Ferner sollten die Materialien nicht die gleichen Informationen beinhalten (nicht redundant sein), sondern sich gegenseitig ergänzen. Darüber hinaus sollten keine interessanten aber irrelevanten Informationen dargeboten werden (coherent). Das Kohärenz-Prinzip spielt bei Horz und Schnotz eine Doppelrolle. Auf der Seite des Lower Order Processing reduziert es die Masse an Information pro Sinneskanal. Auf der Seite des Higher Order Processing bewahrt es davor, verwirrende Informationen zu integrieren.

Eindeutig zuordnungsbar Anweisungen zur Erleichterung von Lernprozessen mit hoher Verarbeitungstiefe in Bezug auf Text und Bild sind das Segmenting und Signaling. Das Segmentieren von Informationen hat das Ziel, komplexe Inhalte mit vielen relevanten Elementen in Einheiten zu packen, die das Arbeitsgedächtnis nicht überfordern und die schrittweise Integration ermöglichen. Hierbei besteht die Kunst nicht nur in der logischen Aufteilung und Aneinanderreihung der Inhalte, sondern auch darin, dass es zu keinen medialen Brüchen kommt, die ihrerseits wieder das Lower Order Processing erschweren.

Als dritten Punkt fordern Horz und Schnotz ebenda die Berücksichtigung individueller Unterschiede. Wie bereits erwähnt lassen sich nicht alle Anweisungen und Prinzipien auf alle Lernenden übertragen. In Abhängigkeit der Fähigkeiten und Fertigkeiten von Lernenden können gut gemeinte Maßnahmen nicht nur am Lehrziel vorbeiführen, sondern auch das Lernen behindern. Aus diesem Grund postulieren Horz und Schnotz

das Prinzip des individuellen Unterschiedes, welches explizit das Vorwissen und die kognitiven Fähigkeiten der einzelnen Lernenden in den Fokus nimmt. Allerdings fehlt es bisher sowohl an Techniken über Repräsentationen, die die Integration von Informationen mit hoher Verarbeitungstiefe fördern, als auch an Forschungsergebnissen zu Auswirkung von multimedialen Lehrmedien auf Lernende mit unterschiedlichen Eigenschaften (Horz und Schnotz 2010, S. 249).

3.3.4 Externe Repräsentationen (Medien) im Biologieunterricht

Neben den angeführten Merkmalen der bisher dargestellten Lerntheorien müssen Repräsentationen aus schulpraktischer Sicht weiteren Anforderungen genügen. Hierbei geht es nicht um deren Beschaffenheit in Design und Ablauf, sondern vor allem um den Zweck, den sie im Unterricht erfüllen sollen. Die Biologie bietet hierbei im Vergleich zu andern Fächern eine hohe Anzahl an Möglichkeiten unterschiedlichster Repräsentationen für den Unterricht und deren vielfältige Kombination ist geradezu charakteristisch für das Fach (Schwanewedel und Nitz 2013, S. 11).

Diese Vielfalt ist ein Segen und ein Fluch zugleich. Schon zur Auswahl des richtigen Formates muss gesagt werden, dass es die beste Repräsentationsform für den Unterricht nicht gibt, genauso wie es die beste Unterrichtsform nicht gibt. Ferner existiert auch keine eindeutige und allumfassende Einteilung aller möglichen Repräsentationsformate, um dieser Vielfalt gerecht zu werden (Köhler 2010). Je nach Klassifikation lassen sich Repräsentationen danach sortieren, ob sie gedruckt sind oder nicht, originale Begegnung zulassen oder in wie weit sie den betreffenden Inhalt abstrahieren oder eben nicht.

Nichtsdestotrotz fordert die Medienkombination im Biologieunterricht den Schülern eine Leistung im Umgang mit zwangsläufig multimodalem Unterricht ab, setzt gleichzeitig aber auch eine Sensibilität der Lehrkräfte im Umgang mit Repräsentationen voraus. Aus diesem Grund müssen sich Lehrkräfte über deren Wirkungsweisen bewusst sein, wenn sie diese zweckdienlich im Unterricht einsetzen möchten. Die Darstellung des produktiven Einsatzes von Medien im Biologieunterricht geht allerdings über den Rahmen dieser Arbeit hinaus und kann in fachdidaktischen Lehrwerken nachgelesen werden (Schaal und Weitzel 2012, Spörhase 2012a, Killermann et al. 2005).

3.3.5 Zusammenfassung

Ausgehend von der Cognitive Load Theory (siehe Seite 37 ff.), ist die Aufnahme und die Verarbeitung von neuen Informationen ein aktiver und hoch selektiver Prozess (Chandler und Sweller 1991). Um bei der begrenzten Kapazität des Gehirns eine hohe Lernleistung zu erzielen, müssen im Unterricht sowohl die zu vermittelnden Inhalte als auch das Lehrmaterial durchdacht und strukturiert sein. Diese Strukturierung kann zum Beispiel durch Auswahl, Reduktion, Markierungen oder Sequenzierung geschehen (siehe Seite 39 ff.). Aus Mayers Theorie zum Lernen mit Medien (2001) lassen sich direkt Anweisungen für den Einsatz von Repräsentationen im Unterricht ableiten. Um Redundanzen zu vermeiden, sollten sich Text und Bild nicht überschneiden, sondern ergänzen. Darüber hinaus sollten sie räumlich nahe beieinander liegen oder zeitlich kurz hintereinander präsentiert werden. Ferner sollten Lehrinhalte in sinnvolle Segmente zerlegt und in nachvollziehbarer Reihenfolge dargeboten werden. Gerade bei unerfahrenen Schülern sollte die Wahrnehmung durch zusätzliche Hinweise oder Farbwahl bei der Erarbeitung gelenkt werden (Holsanova et al. 2009). Diese Prinzipien sind allgemein gültig, geben jedoch keine Hinweise zur individuellen Förderung.

Aus diesem Grund greifen Horz und Schnotz (2010) Mayers Prinzipien auf und ergänzen diese durch das Prinzip des individuellen Unterschieds bei Lernenden in Bezug auf die Verwendung von unterschiedlichen Medien (siehe Seite 41 ff.). Das Prinzip des individuellen Unterschieds besagt, dass über die manipulierbaren Eigenschaften von Medien hinaus immer das Individuum berücksichtigt werden muss. Hierbei gilt es, bei Lernenden vor allem auf unterschiedliches Vorwissen und Vorerfahrungen zu achten. Gerade bei ungeübten Lernern müssen Lehrkräfte mehr Arbeit in die Konstruktion von Lehrmaterialien investieren. Hierbei gilt es, die Materialien so nachvollziehbar wie möglich zu strukturieren, damit den Schülern der Zugang zu den Inhalten erleichtert wird (Kalyuga 2007). Bei der didaktischen Überarbeitung der Lehrmaterialien für leistungsschwächere Schüler sollte allerdings vom gutgemeinten Einsatz von Verzierungen und Schmuckelementen abgesehen werden, da sie die extrinsische Belastungen erhöhen und damit zu weniger Lernleistung führen (Wiegmann et al. 1992).

Als Konsequenz aus der theoriebasierten Rahmung konnten in Bezug auf die Planung und Durchführung dieser Arbeit folgende Punkte abgeleitet werden:

- Die zu testenden Materialien mussten den geltenden Prinzipien der Mediengestaltung und Inhaltsvermittlung entsprechen.
- Die Materialien mussten in ihrer Qualität miteinander vergleichbar sein sowie den gleichen Inhalt transportieren.

- Trotz der Unterschiedlichkeit der beiden Formen des Inputs mussten die Begleitmaterialien (Abbildungen und Hilfen) sowohl mit dem Text als auch mit der Experten-Concept-Map kongruent und einsetzbar sein und durften nicht zu Brüchen im Ablauf der Interviews führen oder eine Form des Inputs bevorzugen.
- Da bei den relevanten Unterschieden zwischen den Schülern vor allem das Vorwissen und die Vorerfahrung eine Rolle spielen, durften die Schüler kaum Wissen zum Thema besitzen.

3.4 Concept-Maps als Lerngelegenheiten für den Unterricht

Entsprechend der Ansprüche der Lehrkräfte aus der Fortbildung (siehe Punkt 4.3.2 ab Seite 99), ist die Entscheidung auf Experten-Concept-Maps als Repräsentationsform zur Differenzierung gefallen. Um diese zu verstehen, wird im folgenden Abschnitt das ursprüngliche Concept-Mapping beschrieben. Hieraus wird dann die Sonderform der Experten-Concept-Map abgeleitet und erklärt, sowie relevante Forschungsergebnisse dazu aufgeführt. Dies wird vertieft durch Forschungsergebnisse, die Lernwirksamkeit von Texten und Experten-Concept-Maps gegenüberstellen. Ferner werden theoriebasierte Ansätze für innere Differenzierung durch Experten-Concept-Maps dargestellt, sowie in einem Exkurs konkrete Handlungsanweisungen für den praktischen Einsatz von Concept-Maps im Schulalltag gegeben.

Einführung: Mind Mapping ist bereits vielen bekannt und wird schon weitläufig im Schulalltag eingesetzt (Davies 2011). Es dient dem Brainstorming, zur Veranschaulichung von Inhalten und der Ergebnissicherung. Concept-Mapping ist eine Methode, die einen Schritt weiter geht. Durch das Concept-Mapping wird Vorwissen aktiviert und Übersichtlichkeit geschaffen, ferner werden Zusammenhänge verdeutlicht und somit das Verständnis von komplexen Themen gefördert (Dohnicht 2004, S. 157). Des Weiteren bietet Concept-Mapping mehr Chancen, vom additiven zum kumulativen Lernen zu kommen. Es ermöglicht eine ganzheitliche Herangehensweise an vielschichtige Themen und Inhalte. Tab. 6 verdeutlicht die Abgrenzung zwischen Mind- und Concept-Maps.

Tab. 6: Abgrenzung zwischen Mind-Map und Concept-Map (in Anlehnung an Eppler 2006, S.203).

	Mind-Map	Concept-Map
Definition	Eine Mind-Map ist ein mehrfarbiges und bildliches Kreisdiagramm, welches semantische oder andere Verbindungen zwischen Teilen des Lernstoffs hierarchisch darstellt	Eine Concept-Map stellt die Beziehungen zwischen Begriffen (Konzepten) dar
Hauptfunktion und Vorteil	Zeigt die Unterthemen einer Wissensdomäne auf individuelle, kreative Art und Weise	Zeigt semantische Beziehungen zwischen den Sub-Konzepten in Beziehung zu einem Hauptgedanken
Typischer Kontext für die Anwendung	Brainstorming, persönliche Notizen, Wiederholung	Unterricht im Klassenzimmer, Selbststudium, Wiederholung

Begriffsbestimmung: Concept-Maps sind graphische Darstellungen von Begriffen und deren Beziehungen untereinander unter Berücksichtigung einer *focus question* (Novak und Cañas 2006). Sie dienen der Organisation und Abbildung von Wissens- und Verstehensstrukturen. Begriffe werden dabei normalerweise in Kästchen und deren Beziehungen untereinander mit Pfeilen dargestellt. Diese Pfeile sind beschriftet, um die Beziehungen der Begriffe untereinander zu verdeutlichen. Die Beschriftung soll kurz und prägnant sein, zum Beispiel: *hat Auswirkung auf*, *teilt sich in*, *frisst* und so weiter. Durch die Vernetzung der wichtigsten Begriffe und die Verknüpfung nebeneinanderstehender Konzepte eignen sich Concept-Maps ganz besonders für den Einsatz in der Biologie (Kinchin 2000, S. 62, Nesbit und Adesope 2006, S. 418, Weidenmann 2006, S. 444, Haugwitz und Sandmann 2009, S. 91). Für Concept-Maps werden viele Synonyme verwendet (Struktur-Lege-Technik, Begriffsnetz, Begriffslandkarte, graphic organizer, knowledge map, node-link map oder auch novakian map). Im Folgenden wird aus Gründen der Lesbarkeit nur noch Concept-Mapping verwendet.

Entstehungsgeschichte: Concept-Mapping wurde in den 60er Jahren von Joseph D. Novak und seinen Mitarbeitern an der Cornell University (New York) entwickelt. Es sollte den Lernzuwachs der Schüler sichtbar machen sowie ein bedeutsames und sinnvolles Lernen ermöglichen. Die Idee des Concept-Mapping basiert auf der Theorie des Konstruktivismus, die besagt, dass Lernende ihr Wissen aktiv und individuell konstruieren. Besonderen Wert legte Novak auf Ausubels Theorie der Erkenntnisgewinnung, die ausdrückt, dass das Vorwissen eines Lernenden große Auswirkung auf das Lernen hat. Folgt man dieser Theorie, muss das Konsequenzen auf die Konzeption von Unterricht haben, denn gerade hier entfaltet sich der Nutzen der Concept-Maps (Novak und Cañas 2006). Diese soll im Folgenden näher betrachtet werden.

Funktion und Nutzen von Concept-Maps: Die graphische Darstellung von Inhalten und Zusammenhängen hat nach Brüning und Saum (Brüning und Saum 2007) mehrere Vorteile:

- Visuelle Symbole werden schneller erkannt und begriffen.
- Eine Concept-Map ist nicht an die Linearität eines Textes gebunden.
- Der minimale Einsatz von Wort und Text dient der Überschaubarkeit und macht das Suchen und Erkennen von Begriffen und Zusammenhängen einfacher.
- Die visuelle Darstellungsform erlaubt ein ganzheitlicheres Verständnis.

Der allgemeine Einsatz von Concept-Maps wird von Plotnick (Plotnick 1997) zusätzlich in vier große Bereiche gegliedert:

- **Kreativität:** Die Erstellung einer Concept-Map ist zu Beginn wie ein Brainstorming. Ideen werden zuerst unkommentiert niedergeschrieben und dann überdacht. Die Neustrukturierung verschafft einen Überblick und kann zu neuen Ideen führen, die in die bisherige Vorstellung integriert werden können.
- **Kommunikation:** Erstellt eine Person eine Concept-Map zu einem Thema bildet dies das jeweilige Wissen einer Person zu einem betreffenden Thema ab. Erstellt eine andere Person eine Concept-Map zum gleichen Thema, werden sich die Ergebnisse unterscheiden. Dieser Unterschied ist fruchtbarer Boden für fachliche Diskussionen. Ein weiterer kommunikativer Punkt ist die gemeinschaftliche Erstellung einer Concept-Map. Hier sind viele Meinungen auf einen Nenner zu bringen, was ohne Diskussion nicht möglich ist und die fachsprachliche Auseinandersetzung mit dem Thema fördert.
- **Lernen:** Aus konstruktivistischer Sicht kann Lernen nur dann sinnvoll stattfinden, wenn es mit Bekanntem in Verbindung gebracht und verknüpft wird. Die Einbettung von Neuem in bereits Gelerntem scheint zudem positive Auswirkungen auf eine längere und nachhaltigere Behaltensleistung zu haben (Tsien 2007).
- **Testen:** Das Anfertigen einer Concept-Map erfordert Wissen und Verstehen. Wenn jemand eigenständig eine Concept-Map zu einem Thema erstellen soll und den Sachverhalt nicht verstanden hat, wird dies schnell deutlich, was dann auch bewertet werden kann. Des Weiteren können Fehlvorstellungen (misconceptions) sichtbar gemacht und bearbeitet werden.

Anfertigung: Novak (Novak und Cañas 2006) gibt drei Voraussetzungen, nach denen Lernende eine Concept-Map erstellen können: Die Inhalte müssen klar und logisch zu begreifen sein und bereits einen Zusammenhang haben, die Lernenden müssen zu dem

Thema Vorwissen besitzen und die Inhalte sollten für die Lernenden in irgendeiner Weise bedeutungsvoll sein.

Des Weiteren sollte eine gelungene Concept-Map eine hierarchische Struktur haben. Novak begründet dies durch Ergebnisse neuerer Forschungen zur Behaltensleistung unseres Gehirns, welches neue Informationen im Kurz-, Arbeits- und Langzeitgedächtnis ähnlich verarbeitet, strukturiert und archiviert und damit leicht abrufbar macht (Tsien 2007).

Die Erstellung und Bearbeitung einer Concept-Map in einem Kontext und unter Berücksichtigung einer *focus question* ist Novak besonders wichtig. Er sagt, dass keine Concept-Map ohne eine präzise Frage oder genau definierte Anschauung begonnen werden sollte, da man sich sonst in Belanglosigkeiten verläuft und den Blick auf das Wesentliche verliert. Man erkennt das daran, dass zu viele Begriffe genannt werden und die Concept-Map damit unübersichtlich wird. Besonders betont er auch, dass eine Concept-Map niemals wirklich fertig ist und immer wieder einer Überarbeitung bedarf. Das liegt daran, dass sich die mentalen Strukturen ändern, wenn man lernt (Novak und Cañas 2006). Veränderungen der mentalen Strukturen spiegeln sich in Unzufriedenheit mit der erstellten Concept-Map wider, die dann revidiert werden muss. Würde die selbst erstellte Concept-Map von einer anderen Person kommentiert oder verändert werden, führte dies zwangsläufig zu einer Diskussion über fachliche Inhalte, was seinerseits eine intensive Auseinandersetzung mit dem Thema fördert.

Das Erstellen einer Concept-Map gliedert sich nach White (White 2002) in verschiedene Phasen, die in Tab. 7 dargestellt sind.

Tab. 7: Phasen der Erstellung einer Concept-Map.

Brainstorming: Hierfür benötigt man das eigene (Vor-)Wissen, Notizen, Lehr- und Fachbücher, das Internet oder andere Quellen. Es geht darum, wichtige Fakten, Namen, Begriffe, Ideen zu einem Thema oder einer Fragestellung zu sammeln. Die Begriffe werden einzeln auf kleinen Kärtchen, Post-Its® oder ähnlichem gesammelt, wenn möglich aber gleich mit einem entsprechenden Computer-Programm erstellt. Ziel ist die größtmögliche Anzahl von Begriffen zu generieren.

Organisation: Nun gilt es, das Gesammelte in eine grobe Struktur zu bringen. Das Bilden von Gruppen und Untergruppen verwandter Begriffe ist zweckdienlich, auch eine Hierarchisierung bietet sich an. Bei der Verschaffung dieses groben Überblicks kommen eventuell noch Begriffe, die während des Brainstormings übersehen wurden und nun noch ergänzt werden können. Auch werden Doppelnennungen oder Synonyme deutlich, die entfernt oder zusammengefasst werden können. Wenn manche Begriffe zu mehreren Gruppen passen könnten ist das nicht tragisch. Diese Begriffe werden in der Verbindungsphase wichtig.

Aufbauskizze: Falls nicht mit einem Computer-Programm gearbeitet wurde sondern mit Kärtchen, wird nun ein großes Blatt Papier, Plakat, Pinnwand oder ähnliches benötigt. Nun gilt es, die Gruppen und Untergruppen zu arrangieren. Hierbei wird hierarchisch vorgegangen. Die wichtigeren Begriffe sind mittig und oben anzuordnen. Das wird soweit fortgeführt, bis die weniger wichtigen Begriffe oder Detailwissen am Rande und unten liegen. Auch sollte darauf geachtet werden, dass ähnliche Konzepte und die Begriffe der Untergruppen nah beieinander liegen, was die Verbindung der Begriffe später übersichtlicher macht.

Verbindungen: Nun werden die Begriffe mit Pfeilen miteinander verbunden. Diese Verbindungspfeile sollten mit Wörtern oder Kurzsätzen versehen werden. Das dient der Spezifizierung der Beziehungen der Begriffe untereinander. Begriffe, die in mehrere Kategorien oder Gruppen passen, verdeutlichen hier, dass nicht nur Begriffe untereinander verbunden sein können, sondern auch ganze Gruppen.

Fertigstellung: In jeder Phase ist das Umordnen, Neustrukturieren, Ergänzen oder Entfernen von Begriffen oder Verbindungen möglich. Wenn keine größeren Änderungen mehr zu vollziehen sind und die eigentliche Concept-Map erstellt ist, geht es in die Phase der Fertigstellung. Vor der Feinarbeit sollten folgende Fragen kontrolliert werden: Sind alle Pfeile und Beschriftungen korrekt? Fehlen noch wichtige Begriffe? Ist der Aufbau logisch und nachvollziehbar? Kann man die Concept-Maps ordentlicher gestalten?

Aufbereitung: In der darauffolgenden Feinarbeitsphase geht es darum, das Thema mit graphischen Mitteln zu unterstützen. Folgende Mittel lassen sich schnell und einfach verwirklichen: Einsatz von Farben, Rahmen, Kästchen, die Form der Pfeile, Buchstabengröße, ergänzende Bilder, Diagramme und ähnliches.

3.4.1 Lernen mit Concept-Maps

Der Einsatz des Concept-Mapping Verfahrens als Methode entspringt dem Konstruktivismus und entspricht im schulischen Setting der konstruktivistischen Didaktik. Das Lernen wird hierbei als selbstgesteuerter Prozess der Wirklichkeits- und Sinnkonstruktion verstanden (Novak und Cañas 2006). Da sich Lernen in jedem Menschen individuell vollzieht, ist die Vorhersage von Lernzuwachs durch Wenn-Dann-Beziehungen relativ ungenau. Daher sollten Lernumgebungen reichhaltige, multimodale und kommunikationsorientierte Anreize schaffen, welche das subjektive Empfinden ansprechen und individuelles Lernen ermöglichen (Siebert 1998). Durch die freie Anordnung, Zu- und Wegnahme von Begriffen oder Verbindungen hat das Concept-Mapping großes Potenzial für das sinnvolle Lernen (Kinchin et al. 2000, S. 44). So können Lernende aufgrund der übersichtlichen Struktur ihr eigenes Wissen ordnen und darstellen. Sie können beim Lesen oder Konstruieren von eigenen Concept-Maps neue Wissensstrukturen durchdringen und in ihr persönliches Wissen integrieren (Haugwitz und Sandmann 2009, S. 90). Ferner scheint Concept-Mapping im Vorteil gegenüber anderen Aktivitäten zu sein wie zum Beispiel dem Schreiben von Zusammenfassungen und Gliederungen (Nesbit und Adesope 2006, S. 434). Des Weiteren ist der Lernzuwachs im Vergleich zu anderen Methoden direkt in den von Schülern erstellten Concept-Maps sichtbar. So zeichnen sich Concept-Maps von fortgeschrittenen Schülern im Vergleich zu Anfängern durch eine höhere Verknüpfungsdichte aus, welche sich nicht nur in der korrekten Anzahl der Verbindungspfeile, sondern auch in der Qualität der Aussage der Pfeilbeschriftungen widerspiegelt (Kinchin 2000, S. 65). Concept-Mapping ist ein effektives Werkzeug zur Darstellung von persönlichem Wissen, allerdings konnte bisher noch kein Vorteil bei der Vermittlung von anwendungsbezogenem Wissen oder der Steigerung der Schülerleistung bei Transferaufgaben nachgewiesen werden (Stern et al. 2003, S. 191, Nesbit und Adesope 2006, S. 434, Haugwitz und Sandmann 2009, S. 102).

Vor- und Nachteile von Concept-Maps

Folgende Vorteile werden in der Literatur diskutiert:

- 1) Die Platzierung von Verknüpfungen führt zu einer niedrigeren kognitiven Belastung des visuellen Gedächtnisses als bei einem Text (siehe 3.3 Seite 36 ff.). Das gilt sowohl für das Suchen von Informationen als auch das Verbinden und Einordnen von ähnlichen Begriffen oder Konzepten (Toth et al. 2002, S. 264, Nesbit und Adesope 2006, S. 418), was durch die direkte Begriff-Relation-Begriff Struktur von Concept-Maps gefördert wird (Haugwitz und Sandmann 2009, S. 103).

- 2) Das Erstellen einer Concept-Map ist im Vergleich zu einem Text weniger routiniert, was die Chance einer metakognitive Auseinandersetzung mit dem Inhalt (Hauser et al. 2006).
- 3) Concept-Mapping in Bezug auf die verbale Codierung scheint effektiver als das Formulieren von Texten zu sein, da dies kein detailliertes und damit zeitaufwendiges Schreiben erfordert, was diese Methode besonders effektiv für kooperative Phasen macht (O'Donnell et al. 2002, Nesbit und Adesope 2006, S. 418–420).
- 4) Die Reduktion von Text und Grammatik geht mit einem weiteren Vorteil von Concept-Mapping zu Text einher. So profitieren vor allem sprachschwächere Schüler von dieser Methode (Nesbit und Adesope 2006, S. 420, Haugwitz und Sandmann 2009, S. 92).

Folgende Nachteile von Concept-Mapping im Vergleich zu Text werden in der Literatur diskutiert:

- 1) Bei einer Concept-Map fehlt eine explizite Leserichtung (Nesbit und Adesope 2006, S. 419), was zum einen den Einstieg in eine Concept-Map erschwert und zum anderen nicht garantiert, dass gerade bei komplexeren Maps auch alles gelesen wird.
- 2) Die Erstellung einer Concept-Map ohne Computer ist sehr aufwendig. Weiterhin sind nachträgliche Änderungen an einer mit Papier und Bleistift erstellte Concept-Map mit großem Mehraufwand verbunden (Chang et al. 2001, S. 22).
- 3) Aufgrund der Individualität und Komplexität von Concept-Maps können diese ein sehr eigenwilliges Design bekommen. Solch eine Concept-Map kann auf Schüler überwältigend und demotivierend wirken (Davies 2011, S. 285).
- 4) Zu Beginn kann es bei Schülern zu einer verminderten Lernleistung kommen, da das Einüben einer neuen Methode mentale Ressourcen benötigt, die beim Lernen fehlen (Hauser et al. 2006, S. 243).

3.4.2 Experten-Concept-Maps, eine Sonderform

Im Bereich der Concept-Maps stellen Experten-Concept-Maps eine Untergruppe dar. Im Gegensatz zu den von Schülern selbst erstellten Concept-Maps werden Experten-Concept-Maps von Lehrkräften (also Experten) erstellt und anschließend den Schülern als Lernwerkzeug zu Verfügung gestellt (Wiegmann et al. 1992, Kinchin et al. 2000, Haugwitz und Sandmann 2009). In der internationalen Forschungsliteratur werden Experten-Concept-Maps auch als *worked-out maps* (O'Donnell et al. 2002, Hauser et al. 2006, Hilbert und Renkl 2008) und neuerdings auch als *instructor-provided concept-maps* bezeichnet (Chou 2013). Genauso wie bei den Concept-Maps liegt der hauptsächliche Nutzen von Experten-Concept-Maps in der Betonung der Zusammen-

hänge in einem Themengebiet und besonders in der Verringerung der kognitiven Belastung durch das Lehrmaterial (siehe Cognitive-Load-Theory in Kapitel 3.3.1 ab Seite 37).

3.4.3 Erwiesene Lernwirksamkeit von Experten-Concept-Maps

Die Struktur der Map beeinflusst das Lernen. Bei dem Vergleich des Aufbaus von unterschiedlichen Experten-Concept-Maps konnte gezeigt werden, dass mit Experten-Concept-Maps, die den Gestalt-Prinzipien wie Symmetrie und Nähe der zusammengehörigen Begriffe folgten (siehe Koffka 1999), größere Lernerfolge bei Lernenden erzielt wurden, als durch netzartig strukturierte (Wiegmann et al. 1992, S. 144). Ferner konnte gezeigt werden, dass schon kleine Veränderung der Eigenschaften einer Experten-Concept-Map unterschiedliche Auswirkungen auf Schüler haben können. Beispielweise scheint die Staffelung von Inhalten über mehrere Concept-Maps keine Auswirkung auf Schüler im Bereich der räumlichen Orientierung zu haben, wohingegen Schüler mit einer geringeren räumlichen Fähigkeit überfordert waren. Auch jedwede Art von Verzierungen (Beschriftungen, unterschiedliche Pfeil- und Linienformen) der Experten-Concept-Map hatten unterschiedliche Wirkungen. So minderten die Ausschmückungen den Lernertrag bei Schülern mit niedrigen verbalen Fähigkeiten, wohingegen sie die Lernleistung bei Schülern mit hohen verbalen Fähigkeiten erhöhten (Wiegmann et al. 1992, S. 148–152).

Bei einer Studie, die die Lernwirksamkeit von komplett selbst erstellten Concept-Maps (construct-by-self) im Vergleich zu vorgefertigten, dafür aber lückenhaften Experten-Concept-Maps (construct-on-scaffold) untersucht, zeigt sich, dass die lückenhaften Experten-Concept-Maps bessere Effekte auf das Lernen in Biologie hat (Chang et al. 2001, S. 21).

In Bezug auf graphische Repräsentationen konnte darüber hinaus gezeigt werden, dass sich das sanfte Leiten von Aufmerksamkeit durch ein Lernmaterial besser auf die Lernwirksamkeit auswirkt als eine komplett freie Herangehensweise der Lernenden an das Material (Holsanova et al. 2009, S. 1224). Für Experten-Concept-Maps bedeutet das, dass sich das Geben eines Startpunktes als lernförderlich auswirkt.

3.4.4 Experten-Concept-Map im Vergleich zu Text

Generell können Experten-Concept-Maps unter bestimmten Umständen als lernwirksamere Alternative zum traditionellen Text angesehen werden (Wiegmann et al. 1992, S. 136). Sowohl im affektiven als auch im kognitiven Bereich scheinen Schüler sich an mehr zentrale Gedanken zu erinnern, wenn sie anstelle eines Textes eine Knowledge-Map verwenden (Hall und O'Donnell 1996, S. 100, O'Donnell et al. 2002, S. 81).

Knowledge-Maps ähneln Experten-Concept-Maps, allerdings unterscheiden sie sich durch die absichtliche Verwendung gemeinsamer Sätze von markierten Links, die darauf abzielen, Leitideen miteinander zu verbinden und nicht Inhalte.

Des Weiteren zeigten Personen mit Experten-Concept-Maps signifikant größere Motivation und Konzentration als Personen, die Text verwendeten (O'Donnell et al. 2002). In einer anderen Studie schätzten die Probanden die Arbeit mit Experten-Concept-Map effektiver ein als die Arbeit mit Text (Patterson et al. 1993). Die instruktionale Wirksamkeit von Experten-Concept-Maps auf den Wissenserwerb von Lernenden konnte auch bei einer Studie in einer Online-Hypertext-Umgebungen nachgewiesen werden (Chou 2013).

3.4.5 Experten-Concept-Maps als Differenzierungsmethode

Bei einer Studie, die Aufbereitung von Lernmaterial in Bezug auf die Lernerfahrung (das Bildungsniveau) untersuchte, konnte gezeigt werden, dass Teilnehmer mit niedrigerer Lernerfahrung viel mehr von kohärent aufbereiteten Materialien, also dem expliziten Geben von sinnbildenden Zusammenhängen, profitierten als Teilnehmer mit höherem Bildungsniveau (Gurlitt und Renkl 2008, S. 416). Andere Studien zeigten, dass sich Experten-Concept-Maps als Kommunikationshilfen während der Partnerarbeit als lernförderlich erwiesen. Das gilt sowohl für Personen mit niedrigen verbalen Fähigkeiten als auch für Personen mit hohen verbalen Fähigkeiten (Patterson et al. 1993, S. 9). Aus diesem Grund ist es nicht verwunderlich, dass Experten-Concept-Maps als Kommunikationswerkzeug besser von Lernenden bewertet werden als Text (O'Donnell et al. 2002, S. 81). Die Ergebnisse von Patterson haben darüber hinaus gezeigt, dass sich gerade Lernende mit niedriger verbaler Fähigkeit und geringem Vorwissen an mehr Dinge erinnern konnten, wenn sie eine Experten-Concept-Map anstelle eines Texts als Kommunikationshilfe benutzen. In scharfem Kontrast zu ihrer tatsächlichen Leistung berichteten Personen mit niedriger verbaler Fähigkeit allerdings auch, dass sie den Text als hilfreicher empfunden haben (Patterson et al. 1993, S. 1–8).

3.4.6 Zusammenfassung

Die bisherigen Ergebnisse aus der naturwissenschaftlichen Bildungsforschung sollten schon jetzt Schulpraktiker dazu ermuntern, die Methode des Concept-Mapping in all ihren Facetten in ihren Biologieunterricht zu integrieren (Nesbit und Adesope 2006, S. 435, Haugwitz und Sandmann 2009, S. 104). Concept-Maps haben über die bisher beschriebenen Wirkungen hinaus noch weiteres Potenzial sowohl für den Unterricht als auch für die Lehr-Lern-Forschung (van Zele et al. 2004, S. 1062, Kinchin 2000, S. 65).

Experten-Concept-Maps sollten in der Bildungsforschung auf unterschiedliche Wissensgebiete übertragen, überprüft und kombiniert werden (O'Donnell et al. 2002, S. 82), der *worked-out* Effekt auf Lernende weiter untersucht (Hauser et al. 2006, S. 245) und die Wirkung auf unterschiedliche Lernergruppen genauer betrachtet werden (Hay et al. 2008, S. 308). Hierbei sollten für die Vergleichbarkeit der noch folgenden Studien besonders die individuellen Voraussetzungen der Lernenden beachtet und genau beschrieben werden. Dies gilt für Studien in der Grundschule sowie der Sekundarstufe I und II mit Augenmerk auf Lese- und Sprachschwierigkeiten (Nesbit und Adesope 2006, S. 434–435). Ferner bedingen die komplexen Zusammenhänge zwischen dem Einsatz von Experten-Concept-Maps und den Vorkenntnissen der Lernenden eine ausführlichere Erklärung und weitere Untersuchungen (O'Donnell et al. 2002, S. 80). Hierfür bieten sich unterschiedliche Lern-Domänen, Lernergruppen und vor allem unterschiedliche Aufgabentypen zur Messung der Lernleistung an (Gurlitt und Renkl 2008, S. 417, Haugwitz und Sandmann 2009, S. 93). Darüber hinaus sollten zur Überprüfung der Lerneffektivität mehr individuelle Voraussetzungen der Probanden und die Lernzeitdauer einbezogen werden (Chang et al. 2001, S. 32). Es besteht momentan noch Uneinigkeit in Bezug auf den Zusammenhang zwischen Experten-Concept-Maps, Lernerfolg und der sprachlichen Fähigkeit von Schülern (Wiegmann et al. 1992; O'Donnell et al. 2002, S. 81, Hilbert und Renkl 2008, S. 69).

Ausgehend von den hier beschriebenen Ergebnissen naturwissenschaftlicher Bildungsforschung verfolgt diese Forschungsarbeit folgendes Ziel: die konsequente Beteiligung von Lehrkräften bei der Konzeption, Planung und Durchführung von diesem Projekt. Die Kooperation mit Lehrkräften führt automatisch zu einer starken schulpraktischen Orientierung. Im Vergleich zu anderen Forschungsarbeiten untersucht dieses Projekt darüber hinaus den *worked-out* Effekt von Experten-Concept-Maps im Vergleich zu Text. Die Überprüfung der Wirksamkeit dieses Ansatzes erfolgt ausschließlich auf dem Wissensgebiet der Biologie in Bezug auf Schüler der Sekundarstufe I. Weiterhin wird nicht nur auf die individuellen Voraussetzungen der Schülern geachtet, sondern darüber hinaus auch nach möglichen Unterschieden und Gemeinsamkeiten von Schü-

lergruppen gesucht. Neben der Lernzeitdauer werden mögliche Effekte des Inputs auf Lernende mittels variierenden Aufgaben untersucht, die zur Lösung unterschiedliche Eigenschaften bei Schülern erfordern und somit zu einer präziseren Beschreibung der Schülerleistung führen, als dies durch Multiple-Choice-Aufgaben möglich wäre.

3.5 Aufgaben im Unterricht

Der Einsatz von Aufgaben im Unterricht ermöglicht problemorientiertes und kumulatives Lernen, erfordert eigenverantwortliches Arbeiten von Schülern und sichert Basiswissen (Rademann und Blume 2001, S. 431). Aufgaben dienen dabei nicht nur der Anwendung und Überprüfung von Wissen, sondern fördern vor allem auch das Verständnis (Neuweg 2008, S. 90). Aufgaben bilden im Unterricht die Schnittstelle zwischen bereits erworbenen Kenntnissen und Fähigkeiten und neu zu erwerbendem Wissen und Können. Dadurch nehmen sie eine Schlüsselposition im Lernprozess der Schüler ein (Knoll 1998, S. 47). So ist es nicht verwunderlich, dass etwa zwei Drittel des Biologieunterrichts (68 %) aus Aufgaben besteht (Jatzwauk et al. 2008, S. 272).

Eine Definitionsbestimmung von Aufgaben im schulischen Kontext

Unter Aufgaben im schulischen Kontext werden Anforderungen verstanden, womit Schüler im Unterricht konfrontiert werden (Blömeke et al. 2006, S. 331). Als pragmatische Definition für eine Aufgabe verwendet Kuhn das Wort *Problemraum*. Dieser Problemraum wird durch die Ausgangssituation zur Bearbeitung, dem Lösungsweg und dem Ergebnis definiert (Kuhn und Müller 2007, S. 148).

Bruder legt bei ihrer Definition mehr Wert auf den Aufforderungsaspekt und beschreibt Aufgaben als eine Ausgangslage mit einer Endsituation, wobei die Transformation der Ausgangs- in die Endsituation dem Lösungsweg entspricht (Bruder 2003, S. 13).

Seel wiederum zerlegt eine Aufgabe in vier Komponenten:

- 1) Die Informationskomponente, in der kurz und prägnant die Aufgabeninhalte dargestellt werden.
- 2) Die Reizkomponente, die eine genaue Frage oder Aufforderung stellt und eine konkret beschreibbare Operation beinhaltet (zum Beispiel: *Ergänze die Wörter oder Ordne die Bilder zu*).
- 3) Die Reaktionskomponente, in der das Lösungsverhalten gezeigt wird und schließlich
- 4) die Rückmeldungskomponente, in der ein Feedback über das gezeigte Lösungsverhalten erfolgt (Seel 1981).

Narciss erweitert die Sichtweise auf Aufgaben, indem sie die Inhalts- und Kognitionskomponente hinzufügt. So sollte Lehrkräften bei der Konstruktion und Bewertung von Aufgaben möglichst immer alle lösungsnotwendigen kognitiven Operationen bewusst

sein. Darüber hinaus spielt die Art und Weise, wie Aufgaben den Schülern präsentiert werden, eine wichtige Rolle (Narciss und Proske 2001, S. 5).

Dieser Arbeit wird die Definition von Aufgaben durch Bruder (2003) zugrunde gelegt. Ab Seite 60 werden Aufgaben anhand von konkreten Beispielen genauer dargestellt und durch zusätzliche Kriterien von Gropengießer (2006) in einen biologiedidaktischen Kontext gestellt.

3.5.1 Formale Aspekte von Aufgaben

Für eine gewinnbringende Diskussion sind formale Aspekte und Begriffe von Nutzen, wonach sich alle Aufgaben beschreiben lassen. Die Begriffe sind vor allem nur dann von Nutzen, wenn sie einheitlich verwendet werden (Thonhauser 2008).

Ein praxisorientiertes Merkmal zur Unterscheidung von Aufgaben ist die Art und Weise, wie die Aufgaben bearbeitet werden können. Das kann mündlich, schriftlich oder fachpraktisch erfolgen (KMK 2004). Zur Verdeutlichung der Wichtigkeit dieses Merkmals dient das Baden-Württembergische Schulgesetz. Dieses fordert, dass bei der Notenfindung zwischen mündlicher und schriftlicher Leistung unterschieden werden muss. Bei einer mündlichen Aufgabe kann die Zeit zur Beantwortung beispielsweise viel geringer als bei einer schriftlichen Abfrage ausfallen, wohingegen bei einer schriftlichen Aufgabe neben der korrekten Beantwortung auch noch ein erhebliches Maß an Lese- und Schreibfähigkeit hinzukommt. Unter Berücksichtigung der unterschiedlichen Fähigkeiten der Schüler sollte daher auf eine gesunde Mischung zwischen mündlichen und schriftlichen Aufgaben im Unterricht geachtet werden (Kultusministerium Baden-Württemberg).

Eine gesonderte Stellung nehmen die fachpraktischen Leistungen ein. Zwar bietet sich die Fachpraxis im Biologieunterricht an, jedoch haben diese Leistungen bisher kaum Einfluss auf die Note. Das liegt zum einen daran, dass eine fachpraktische Leistungssituation äußerst material- und zeitaufwendig ist und zum anderen daran, dass sich die Testgütekriterien (Reliabilität, Objektivität und Validität) nicht so einfach auf solche Leistungen übertragen lassen.

Die Sozialform ist ein weiteres Kriterium, Aufgaben zu beschreiben (Köhler 2004). Es macht einen Unterschied, ob eine Aufgabe in Einzelarbeit oder kooperativ gelöst werden soll. Auch hier gibt der Zweck, den eine Aufgabe im Unterricht erfüllen soll, deren Sozialform vor. Soll zum Beispiel die Kommunikationsfähigkeit geübt werden, so macht Einzelarbeit wenig Sinn. Soll sich ein Schüler ganz speziell auf einem Gebiet verbessern, ist Teamarbeit nicht das förderlichste, es sei denn, Teamarbeit ist das zu fördernde Ziel.

Ein weiteres Unterscheidungsmerkmal ist das Antwortformat der Aufgabe. Unter Aufgabenformat versteht man die Art und Weise der Aufgabenstellung und der daraus folgenden Aufgabenbeantwortung. Hieraus ergeben sich drei große Formate auf die im Folgenden genauer eingegangen wird (Graf 2001, Porst 2009, Schmiemann 2013):

- Geschlossene (gebundene) Antwortformate sind solche, bei denen Antwortalternativen vorgegeben sind. Dazu gehören: Zuordnungsaufgaben / Umordnungsaufgaben, Einfach- und Mehrfachwahlaufgaben (Multiple-Choice) und Richtig-Falsch-Aufgaben. Geschlossene Aufgabenformate haben den Vorteil der eindeutigen Erfassung und damit der objektiven Auswertung von Ergebnissen. Auch sind sie äußerst ökonomisch hinsichtlich ihrer Durchführung, da die Lösungszeit für eine Aufgabe in der Regel gering ist, womit auch die Anzahl der möglichen Aufgaben steigt, was wiederum eine treffsichere Feststellung der Schülerleistung ermöglicht. Da es nur eine richtige Lösung gibt, steigt auch die Objektivität der Auswertung. Ein Nachteil von Aufgaben mit geschlossenem Antwortformat ist die zeitintensive Konstruktion und die hohe Ratewahrscheinlichkeit. Es muss auch erwähnt werden, dass das Entwickeln und Abprüfen von komplexen Kompetenzen oder Kreativität mit diesem Aufgabenformat nicht zu bewerkstelligen ist.
- Halboffene Antwortformate sind Aufgaben, bei denen die Aufgabe zum Beispiel durch ein Schlüsselwort, ein Symbol oder eine kurze zeichnerische Darstellung vollendet wird. Dazu gehören Lückentexte und Ergänzungsaufgaben. Bei diesem Typ von Aufgaben sind individuelle und etwas freiere Antworten möglich. Die Zeit für die Entwicklung dieser Art von Aufgaben ist geringer und der Zufallseinfluss entfällt gänzlich. Die individuelleren Antworten erschweren eine objektive Auswertung, außerdem steigt die Bearbeitungszeit.
- Offene (freie) Antwortformate erfordern umfangreichere, selbst formulierte Antworten, bei denen vielfältige Lösungswege möglich sind. Dazu gehören: Zeichnungen und Skizzen, Kurzaufsätze, Formulierung von Lösungswegen und Begründungen. Dieser Antworttyp fördert individuelle Lösungen und lässt auch verschiedene Lösungswege zu. Der Zeitaufwand für die Entwicklung offener Aufgaben ist vergleichsweise gering. Die Bearbeitung komplexer Probleme erfordert von den Lernenden Kreativität, vielschichtiges Denken und vor allem Zeit. Die verschiedenen Lösungsmöglichkeiten mindern die Objektivität und erhöhen den Korrekturaufwand der Lehrkraft.

Aus dieser Einteilung wird ersichtlich, dass jedes Aufgabenformat bestimmte Vor- und Nachteile bezüglich der Aufgabenstellung, der Objektivität der Auswertung und der zeitlichen Ökonomie hat. Ferner bezieht sich diese Einteilung auf schriftliche Ant-

worten und lässt sich nicht auf mündliche Aufgaben oder fachpraktische Arbeitsweisen übertragen.

Aufgabentypen und Kombinationen in der Schulpraxis

Basierend auf den bereits erwähnten Definitionen von Aufgaben (siehe Seite 56 ff.) lassen sich Aufgaben in verschiedene Grundtypen unterteilen. Diese Unterteilung basiert auf der unterschiedlichen Zusammensetzung der immer gleichen Grundeinheiten aller Aufgaben.

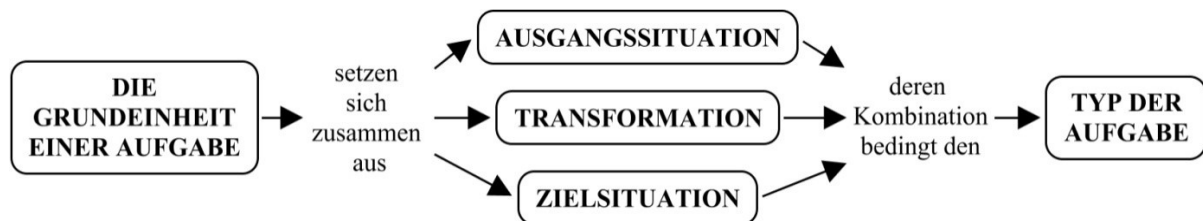


Abb. 13: Der Aufgabentyp in Abhängigkeit zu den Grundeinheiten einer Aufgabe (nach Bruder 2003).

Damit die Bezeichnung *Ausgangssituation* möglichst viele Aufgaben abdecken kann, wurde bewusst der Begriff *Situation* gewählt. Eine Situation kann den Ausgangszustand einer Aufgabe beschreiben, die Darstellungsform von Informationen bedeuten oder etwa die Ausgangslage vor einer Veränderung verdeutlichen. Die zweite Grundeinheit ist die *Transformation*. Hierbei ist jedwede Aktion gemeint, die der Umsetzung von einer Ausgangssituation in eine Zielsituation dient oder umgekehrt. Als dritte Grundeinheit wird die *Zielsituation* dargestellt. Die einfachste Definition einer Zielsituation ist das Ergebnis einer Aufgabe. Dies kann zum einen die Lösung sein, jedoch auch eine andere Form der Darstellung oder Präsentation. Zur Verdeutlichung werden im Folgenden die drei Grundeinheiten erklärt und an Beispielen veranschaulicht.

Die Zuweisung verschiedener Aufgabentypen basiert hierbei auf Kriterien von Bruder (2003) und Gropengießer (2006). Die sich immer wiederholende Tabelle auf der linken Seite fasst den Zustand der Aufgabe zusammen: **A** steht für Ausgangssituation, **T** für Transformation und **Z** für Zielsituation. Das **X** steht für *ist gegeben* und das Minusymbol steht für *ist nicht gegeben*.

A	T	Z
X	X	-

Aufgaben, bei denen die Ausgangssituation (Frage / Aufforderung) und die Transformation gegeben sind, gehören zu den Klassikern. Hierbei handelt es sich um Grundaufgaben, bei denen *nur* die Lösung erarbeitet werden soll. Als Beispiel für diesen Aufgabentyp dient die vereinfachte Fotosynthesegleichung mit der dazugehörigen Arbeitsanweisung:

Lese dir die Gleichung sorgfältig durch und ergänze die Lücken: Unter Anwesenheit von Sonnenlicht reagieren $6 \text{H}_2\text{O} + 6 \text{CO}_2$ in intakten Chloroplasten zu _____ + _____?

Wasser, Kohlenstoffdioxid, Licht und Chloroplasten sind die Ausgangssituation. Sie sollen addiert und korrekt kombiniert werden (Transformation). Schließlich soll das Ergebnis genannt werden (Zielsituation).

A	T	Z
-	X	X

 Die meisten dieser Grundaufgaben lassen sich auch in Umkehraufgaben verwandeln. Folgende Aufgabe wäre möglich:

Formuliere zu dem Experiment in deinem Schulbuch eine passende Forschungsfrage (Gropengießer 2006).

A	T	Z
X	-	-

 Je weniger der drei Aufgabensituationen vorgegeben sind, desto offener wird eine Aufgabe. Das bedeutet für Lehrkräfte mehr Korrekturaufwand, da es meist mehr als eine mögliche Lösung der Aufgabe gibt. Für Lernende heißt das, dass sie sich intensiver mit der Aufgabe beschäftigen müssen, da eigenes Denken gefordert ist.

Hier eine mögliche Ausgangssituation: Die Lernenden haben eine pflanzliche Zelle präpariert und betrachten diese durch das Mikroskop. Die Aufgabe lautet:

(A) Findet einen Weg, die Größe der Pflanzenzellen unter dem Mikroskop zu bestimmen (Transformation). (B) Wendet das von euch entwickelte Verfahren an und bestimmt die Größe der Zellen (Zielsituation).

Die Methoden und Ergebnisse werden gesammelt und verglichen. Die Lernenden werden nach einer Zielsituation gefragt und es wird vorerst nicht gesagt, wie man diese erreichen kann. Dies stellt sie vor ein Problem, wovon sich auch der Typ der Aufgabe ableitet, die Problemaufgabe. Da man mit dem Wort Problemaufgabe jedoch auch einige andere Aufgaben beschreiben könnte, wird dieser Aufgabentyp auch als Methodenfindungs-Aufgabe bezeichnet. Die Lernenden müssen sich den Weg zum Ziel selbst erarbeiten.

A	T	Z
-	-	X

 Beispielhaft für eine Aufgabe, deren Zielzustand gegeben ist, wäre folgende Aufgabenstellung: Ein Zeitungsartikel besagt, dass das Wasser aus einem Gewässer (am besten in der Nähe der Schule) nicht mehr getrunken werden darf. Die Aufgabe könnte lauten:

Wie könnte man das Wasser untersuchen und welche Stoffe können dazu führen, dass man das Wasser nicht mehr trinken darf?

Diese Problemumkehraufgabe ähnelt der Methodenfindungsaufgabe, geht aber in die andere Richtung.

A	T	Z
X	-	X

 Ist sowohl die Ausgangssituation, als auch der Zielzustand einer Aufgabe gegeben, so handelt es sich um eine Begründungsaufgabe. Beispiel:

Wenn man Wasser und Kohlenstoffdioxid mischen und in die Sonne stellen würde, so würde daraus niemals Zucker und Sauerstoffdioxid entstehen. Was ist für eine gelingende Fotosynthese unbedingt erforderlich?

A	T	Z
-	X	-

 Wenn nur die Transformation gegeben ist, müssen Anwendungen von dieser gefunden und in einen Zusammenhang gebracht werden. Bei der geforderten Transformation sollten Vorgänge gewählt werden, die einen generellen Einfluss auf das Thema haben und in verschiedenen Formen immer wiederkehren, zum Beispiel:

Das Prinzip der Oberflächenvergrößerung kommt in der Natur oft vor. Finde drei Beispiele aus der Natur und erkläre, zu was die Oberflächenvergrößerung in dem jeweiligen Fall dient.

Diesen Aufgabentyp nennt man (Methoden-)Anwendungsaufgabe oder auch Anwendungssuche. Weiterhin kann dieser Typ von Aufgabe auch als Transformationsaufgabe bezeichnet werden, wenn das Wissen in unbekanntem Situationen angewendet werden muss. Für die Schulpraxis ist wichtig zu erwähnen, dass diese Aufgabentypen im Unterricht behandelt und geübt werden sollten, bevor sie in Leistungssituationen zur Diagnose von Transferleistungen eingesetzt werden. Denn mehr Wissen bedeutet nicht, dass man es auch in anderen Situationen besser anwenden kann. Aus diesem Grund müssen Transferleistungen geschult und geübt werden (Hamann 2006, S. 92).

A	T	Z
X	X	X

 Als Abschluss oder Einstieg in eine neue Unterrichtseinheit, können auch Beispielaufgaben bearbeitet werden. Beispielaufgaben beinhalten alle drei Grundeinheiten einer Aufgabe und sind somit vollständig gelöst. Anhand dieser Musterlösungen können Schüler eine komplexe Aufgabe nachvollziehen oder für sich ableiten, wie man an eine ähnliche Aufgabe herangehen könnte. Des Weiteren können Beispielaufgaben den bisherigen Lehrstoff wiederholen und gleichzeitig zu einem neuen Unterrichtsthema überleiten. Verdeutlicht wird dies durch die folgende Aufgabenstellung:

Pflanzen benötigen Wasser und Kohlenstoffdioxid zur Fotosynthese. Sie gelingt nur mit Licht und Chlorophyll in intakten Chloroplasten. Das Ergebnis der Fotosynthese ist Glucose und Sauerstoff. Dies ist soweit richtig, aber was machen Pflanzen nachts? Warum brauchen manche Pflanzen Insekten und andere nicht? Was ist eine Symbiose? Warum wachsen in Deutschland keine Palmen?

A	T	Z
-	-	-

Eine komplett offene Aufgabe hat weder eine konkrete Ausgangssituation, eine geforderte Transformation noch einen einzig richtigen Zielzustand. Diese offene Situation macht die Aufgabe äußerst komplex. Dies gilt sowohl für die Bearbeitung einer Aufgabe durch Lernende als auch für den Umgang mit den Ergebnissen für Lehrkräfte. Bei diesen Aufgaben ist auf beiden Seiten Kreativität gefordert:

Ein Flug zum Mars dauert mehrere Jahre. Du musst dich um die Verpflegung der Mars-Crew kümmern. Was würdest du mitnehmen, damit sich deine Freunde im Raumschiff über diesen Zeitraum gesund ernähren?

Eine ähnliche Aufgabe wäre die eigenständige Aufgabenentwicklung. Hierbei sollen Lernende zu einem vorgegebenen Thema selbst Aufgaben entwickeln:

Formuliere drei für dich interessante Aufgaben zum Thema Pflanzen, die bisher nicht im Unterricht vorgekommen sind.

Die hier dargestellte Aufteilung in verschiedene Aufgaben-Typen hat zwei Vorteile: Zum einen lassen sich Aufgaben einem konkreten Typ zuordnen. Diese Zuordnung ermöglicht es, Häufungen von nur einem Typ aufzudecken und Monotonie im Aufgabeneinsatz aufzudecken und entsprechend entgegen zu wirken. Zum anderen gibt diese Auflistung indirekt einen Leitfaden zur Aufgabenkonstruktion, denn „mithilfe dieser acht Typen kann etwas gelingen, das aus lernpsychologischer Sicht immer wieder gefordert wird: Ein Wechsel der Blickwinkel und Vernetzungen innerhalb eines Themas“ (Bruder 2003, S. 14). Der Lernzuwachs bei Schülern im Unterricht wird definitiv besser sein, je mehr unterschiedliche Aufgaben zu einem Thema gestellt werden (Hammann 2006).

3.5.2 Funktionen von Aufgaben im Unterricht

Aufgaben verfolgen im schulischen Kontext ein konkretes Ziel auch wenn die Definitionen von Aufgaben in der Literatur auseinander gehen (siehe Kapitel 3.5 auf Seite 56 ff.). Sie dienen der Wissensvermittlung und sollen: „ein intelligent geordnetes, in sich vernetztes, in verschiedenen Situationen erprobtes und flexibel anpassbares Wissen, welches Fakten-, Konzept-, Theorie-, Methoden- und Prozesswissen gleichermaßen beinhaltet vermitteln“ (BLK 1997, S. 15).

Nach der Einführung der Bildungsstandards durch die KMK (2004) wurde in der Bundesrepublik Deutschland von der Input-orientierten Auffassung von Lehrplänen auf Output-orientierte Bildungspläne umgestellt. Hierbei wurden Bildungsstandards festgelegt, die verbindlich definieren, welche Kompetenzen die Schüler in einem Fach bis

zu einem bestimmten Zeitpunkt erworben haben sollen. Bei diesen Kompetenzen handelt es sich nach Weinert um die verfügbaren Fähigkeiten und Fertigkeiten um Probleme zu lösen sowie die damit verbundene Fähigkeit, Problemlösungen in variablen Situationen zu nutzen (Weinert 2001). Die Bildungsstandards für die naturwissenschaftlichen Fächer stimmen hierbei zwar in zahlreichen Formulierungen überein, sind jedoch disziplinspezifisch formuliert. Dies zeigt die Nähe der Fächer untereinander, ohne dabei deren individuellen Charakter aufzugeben (Kauertz et al. 2010, S. 137). Für jedes naturwissenschaftliche Schulfach wurden dementsprechend auch eigene Kompetenzbereiche entworfen. Für die Biologie sind das die Bereiche Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung. Die Kompetenzbereiche sind nicht als getrennte Einheiten zu verstehen, sondern vielmehr als unterschiedliche Blickwinkel auf die jeweiligen Fachinhalte (Kauertz et al. 2010, S. 140, KMK 2005).

Für die Erfüllung des Output-orientierten Bildungsauftrages sind nach Hammann ganz besonders Aufgaben geeignet, denn nur durch Aufgaben werden Schüler mit konkreten Anforderungen konfrontiert, anhand derer sie ihre Kompetenzen zeigen können (Hammann 2006, S. 85). Kompetenzen können nicht unterrichtet werden, Kompetenzen können nur angebahnt und erlebbar gemacht werden, was wiederum nur durch Aufgaben gelingen kann (Rademann und Blume 2001, S. 431). Die Kompetenzorientierung kann somit auch als Schülerorientierung verstanden werden, die die Inhalte in Beziehung zu den Schülern setzt (Leuders 2009, S. 5).

Aus der Kompetenzorientierung ergeben sich nach Hammann noch weitere Möglichkeiten. So ist es in Bezug auf Aufgaben möglich, diese nicht nur nach den bereits erwähnten Oberflächenmerkmalen zu klassifizieren (siehe Kap. 3.5.1 auf Seite 57 ff.), sondern auch nach den betreffenden Kompetenzen, die zur Lösung der Aufgabe notwendig sind. Kompetenzorientierung und Aufgabenvielfalt bedingen sich hierbei gegenseitig, denn Aufgabenvielfalt resultiert aus den verschiedenen Kompetenzen, die anhand von Aufgaben gefördert werden können. Gleichzeitig werden vornehmlich nur die Kompetenzen im Unterricht angebahnt, die durch Aufgaben bedient werden (Hammann 2006, S. 85). Aus diesem Grund stellen Bildungsstandards eine Quelle für die Entwicklung von Aufgaben dar. Für kompetenzbereichsorientierte Aufgaben, nennt er zum Beispiel neben Aufgaben zu Kreativität und divergentem Denken auch Aufgaben zum Beschreiben von naturwissenschaftlichen Phänomenen, zum Erklären und Vorhersagen (Fachwissen), zum Ziehen von Schlussfolgerungen (Bewertung), zum Verstehen von naturwissenschaftlichen Untersuchungen und zum Erkennen von naturwissenschaftlichen Fragestellungen sowie zum Erbringen von Nachweisen und Indizien (Erkenntnisgewinnung). Ferner müssen Schüler ihr Wissen über naturwissenschaftliche Sachverhalte und den daraus resultierenden Meinungen und Entschei-

dungen durch entsprechende Aufgaben mit- und untereinander besprechen (Kommunikation) (Hammann 2006, S. 87).

Der von Hammann erwähnten Weiterentwicklung der Aufgabenkultur wird dementsprechend ein beträchtliches Potenzial zur nachhaltigen Qualitätssteigerung des naturwissenschaftlichen Unterrichts zugeschrieben (Kühn 2011, S. 35, Schaal und Weitzel 2012, S. 132). Abgesehen von ihrer Kompetenz- bzw. Schülerorientierung sind Aufgaben nach Rademann nämlich ein wesentlicher Bestandteil der Planung und Durchführung von Unterricht (Rademann und Blume 2001, S. 431). Sie initiieren dabei nicht nur die intensive Beschäftigung mit biologischen Phänomenen, sondern bestimmen auch die Zielrichtung und das Ergebnis der Lernprozesse (Bayrhuber 2000, S. 210). Neben der Lehrzielorientierung geben komplexere Aufgaben mehr oder weniger explizit bestimmte Abläufe und Methodenarrangements vor und strukturieren gleichzeitig auch das Unterrichtsgeschehen (Jatzwauk et al. 2008, S. 264). Andererseits lassen sich Aufgaben aber auch in Abhängigkeit zum Unterricht entwerfen, denn je nach Position im Unterricht verfolgen Aufgaben unterschiedliche Ziele (Leuders 2009). Sollen Aufgaben zum Beispiel am Anfang einer Unterrichtseinheit in ein Thema einführen, sollten sie eher motivieren und Vorwissen aktivieren. Zur Verdeutlichung sind in Tab. 8 Unterrichtsphasen mit exemplarischen Zielen nach Häußler aufgeführt (Häußler und Lind 2000).

Tab. 8: Unterrichtsphasen und Aufgabenziele in Anlehnung an Häußler (Häußler und Lind 2000).

Phasen	Aufgabenziele
Einstiegsphase	Motivierung, Warm-up, vorstellen der Inhalte und Kontexte die zur Erarbeitung der Aufgabe nötig sind, Vorwissen aktivieren
Erarbeitungsphase	Selbstständiges Arbeiten, Umwandlung einzelner Wissensselemente in anwendungsfähiges Wissen
Übungsphase	Festigung von Gelerntem, Übertragung auf neue Situationen
Wiederholungsphase	Vernetzung des neu Gelernten mit bereits gelerntem Stoff
Individualphase	Anpassung an individuellen Lernfortschritt
Gruppenunterricht	Möglichkeiten zum wechselseitigen Helfen und Lehren und Kommunizieren
Hausaufgaben	Anpassung an den individuellen Lernfortschritt, Einzelarbeit, Wiederholung, Festigung, Anwendung
Prüfungsphase	Lernerfolgsmessung

Abgesehen von den offensichtlichen Lern- und Testfunktionen (siehe Seite 67 ff.) geben Aufgaben Rückmeldungen über den Lehrerfolg an Lehrkräfte oder Lernfortschritte der jeweiligen Schüler im Unterricht. Ferner können Schüler bei Aufgaben erkennen, was die Lehrkraft verlangt und ob sie gut damit zurechtkommen oder nicht. Daraus können sie bis zur nächsten Klassenarbeit ableiten, ob sie sich intensiver mit dem Thema auseinandersetzen sollten oder nicht.

Für Lehrer lassen sich aus Aufgaben gleich mehrere Rückschlüsse ziehen. Aufgaben können Alltagskonzepte, „Fehlvorstellungen“ und träges Wissen bei Schülern auf-

decken, die als Ausgangspunkte für problemorientierten Unterricht verwendet werden können. Zum anderen können durch Aufgaben auch andere Fehlerquellen aufgespürt werden. Das macht Aufgaben zu einem Mittel der Diagnose (Müller und Helmke 2008, S. 40) und dienen Lehrkräften als Evaluationsinstrument des eigenen Unterrichtsertrags (Thonhauser 2008, S. 13). Die Evaluation des eigenen Unterrichts kann entsprechend der Frage erfolgen: *Können die Schüler am Ende meines Unterrichts das, was ich ihnen vermitteln wollte?*

3.5.3 Unterscheidung zwischen Test- und Lernaufgaben

Aufgaben in der Schule lassen sich in Test- und Lernaufgaben unterteilen. Markante Unterschiede werden in Tab. 9 aufgeführt und im Folgenden erläutert.

Tab. 9: Unterschiede zwischen Test- und Lernaufgaben (in Anlehnung an Leisen 2010, S. 11).

Lernaufgaben	Leistungsaufgaben
haben unterschiedliche Niveaus	haben häufig ein einziges Niveau
erlauben individuelle Hilfestellung	erlauben keine Hilfe
stehen untereinander im thematischen Zusammenhang	stehen untereinander häufig nicht im Zusammenhang
fördern integrativ unterschiedliche Kompetenzen	überprüfen meist nur eine Kompetenz
dienen dem individuellen Lernfortschritt	dienen dem Leistungsvergleich
sind vielfältig im Lösungsweg und in der Lösungsdarstellung	haben eine eindeutige Lösung
unterstützen den individuellen Lernprozess	diagnostizieren und stellen den individuellen Förderbedarf fest
brauchen Fehler, um aus Fehlern zu lernen	werden korrigiert (d.h. nur das Richtige wird bewertet)
können durch Schüler in unterschiedlicher Zeit bearbeitet werden	sollen in der gleichen Zeit gelöst werden
erlauben eine große Vielfalt an Methoden	Sind beschränkt auf wenige Antwortformate (meistens schriftlich)

Diese Unterscheidung ist wichtig, da eine bewußte Trennung nicht nur Auswirkungen auf die Aufgaben, sondern auch auf den Unterricht hat. Eine Testaufgabe sollte zum Beispiel objektiv, reliabel und valide sein. Die Lernenden sollen ihr erworbenes Wissen und Können beweisen und dementsprechend benotet werden. Bei einer Lernaufgabe hingegen ist ein Fehler kein Mangel, sondern viel eher ein Zeichen für Fortschritt (Helmke 2006b). Das bezieht sich nicht auf Plattitüden wie *Aus Fehlern wird man klug*, sondern zeigt, dass auf dem Weg zum wissenschaftlich korrekten Wissen und Können Fehler unvermeidlich sind. Der Umgang mit Fehlern macht den Unterschied. Bei einer Testaufgabe wird meist mit einem *richtig* oder *falsch* bewertet, bei einer Lernaufgabe steht dagegen das *warum richtig* oder *warum falsch* im Mittelpunkt. Folglich können psychometrische Gütekriterien von Testaufgaben nicht ohne weiteres auf Lernaufgaben übertragen werden (Schecker und Parchmann 2006, S. 62). Fehler dürfen dabei nicht sanktioniert, sondern als Lernchance gesehen werden (Müller und Helmke 2008, S. 37). Aus diesem Grund müssen Lern- und Leistungssituationen im Unterricht strikt voneinander getrennt und deren Vermischung vermieden werden (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung 1997, S. 25). Anderenfalls versuchen Schüler Fehler im Unterricht zu vermeiden und sagen irgendwann eher nichts, bevor sie etwas Falsches sagen, wofür sie dann eine negative Note bekommen könnten. Die Unklarheit zwischen Lern- und Testsituation im bestehenden Prüfungs- und Benotungssystem hat aber nicht nur Auswirkungen auf das Verhalten von Lernenden, sondern auch auf das *was gelernt* und *wie gelernt* wird. Die

Folge ist das Pauken von zusammenhangslosen Wissens-elementen, dessen standesgemäße Reproduktion in der jeweiligen Testsituation, gefolgt vom großen Vergessen (Winter 2008, S. 120).

Mehr als 75% der Testaufgaben zielen auf die Reproduktion von Wissen ab

Eine Analyse von 600 eingesetzten Testaufgaben im Biologieunterricht hat gezeigt, dass sich die Kultur des zusammenhangslosen Lernens auszahlt (Germ und Harms 2010). So beschränken sich 76,2% der analysierten Testaufgaben auf Wissen. *Verstehen* wird demgegenüber seltener erfasst (16%), höhere kognitive Fähigkeiten sind sichtlich unterrepräsentiert (7,3%) und Aufgaben zum Anforderungsbereich *Bewerten* treten nahezu überhaupt nicht auf (0,5%) (Germ und Harms 2010, S. 13). Ferner sind nach Germ die Kompetenz- und Anforderungsbereiche der nationalen Bildungsstandards in den Testaufgaben im Biologieunterricht nur eingeschränkt umgesetzt. In Anbetracht der Auswertung von Testaufgaben ist die starke Dominanz von offenen Antwortformaten aus zwei Gründen kritisch zu betrachten: Die Studie hat zum einen ergeben, dass über 90% der gestellten Aufgaben Freiantworten in schriftlicher Form verlangen, was in Anbetracht des sehr hohen Anteils an reproduktiven Aufgaben der einfachen Reproduktion von Fachinhalten entspricht. Zu anderen ist bekannt, dass viele Schüler zwar das nötige Fachwissen zur Beantwortung der Aufgabe besitzen, sie allerdings große Probleme haben, dieses ausreichend zu verbalisieren (Neuweg 2008, S. 86). Daraus kann gefolgert werden, dass häufig eher sprachliche Fähigkeiten als Wissen und Können in Biologie überprüft werden.

Bei der Analyse der Abituraufgaben in den naturwissenschaftlichen Fächern zeigt sich das gleiche Bild (Kühn 2011). Abituraufgaben enthalten ausschließlich Aufgabenstellungen mit freiem Antwortformat und zielen überwiegend auf das Abfragen von Wissen ab. Im Vergleich dazu fällt der Anteil anwendungsorientierte Aufgaben marginal aus. Der Anteil offener Aufgabenstellungen ist in allen Fächern gering und fällt unterschiedlich aus. Auf die jeweiligen Fächer verteilt liegt der Prozentanteil von offenen Aufgabenstellungen im Abitur in Physik bei 16,5 %, in Chemie bei 2,9 % in Chemie und in Biologie bei 0,8 % (Kühn 2011, S. 44 f.). Nach Kühn hat es die durch den Bildungsplan bedingte Weiterentwicklung der Aufgabenkultur bisher nicht in die schriftlichen Abiturprüfungen der naturwissenschaftlichen Fächer geschafft. Ferner wurde die Vergleichbarkeit der Prüfungsanforderungen trotz bundesweit einheitlicher Prüfungsstandards in den Naturwissenschaften zwischen den Ländern und Fächern bisher nicht erreicht (Kühn 2011, S. 51).

Aufgaben zum Lernen

Genauso wie Testaufgaben müssen Lernaufgaben valide sein, daher muss bei deren Bearbeitung sichergestellt werden, dass etwas gelernt wird und möglichst auch die Inhalte, die gelernt werden sollen (Leutner et al. 2008, S. 171). Im Gegensatz zu Testaufgaben, stehen Lernaufgaben jedoch am Anfang eines Lehr-Lernprozesses, da sie sowohl formal als auch inhaltlich den Weg im Unterricht weisen (Thonhauser 2008, S. 15). Sofern Lernaufgaben richtig eingesetzt werden, entlasten sie dabei den Unterricht von der bereits erwähnten Aura latenter Überprüfung, welche oftmals während eines fragend-entwickelnden Unterrichts entsteht (Stäudel 2004, S. 5, Jatzwauk et al. 2008, S. 264). Lernaufgaben zielen auf das selbstständige und individuelle Erschließen von Neuem und den Aufbau von persönlichen Wissensstrukturen ab (Leisen 2006, S. 263). Im Unterrichtsgeschehen können Lernaufgaben aber noch andere Ziele als das ausschließliche Neuerlernen verfolgen. So geht es bei Lernaufgaben auch um die Wiederholung von Inhalten, der Anwendung von Wissen, dessen Reaktivierung und Festigung (Jatzwauk et al. 2008). Des Weiteren kann Wissen durch Aufgaben aktualisiert oder bewusst mit bereits Gelerntem verknüpft werden (Jatzwauk et al. 2008, S. 264). Lernaufgaben ermöglichen die Aktivierung von Vorwissen und das gezielte Üben von Transferleistungen (Thonhauser, S. 190).

So unterschiedlich beide Formen von Aufgaben sind, so haben sie doch beide ihre Berechtigung im Unterricht, denn nur an Aufgaben lernen wir und nur über Aufgaben ist Lernerfolg diagnostizierbar (Neuweg 2008, S. 84). Die Qualität des Unterrichts wird nach Leisen also nicht nur durch Präambeln und Bildungsstandards verbessert, sondern vor allem durch gute Aufgaben, die im Klassenzimmer auch ankommen (Leisen 2005, S. 306). Aus diesem Grund werden im folgenden Kapitel die Gütekriterien von Lernaufgaben aufgeführt und in Abb. 14 auf der folgenden Seite die wesentlichen Aspekte von Aufgaben im Unterricht zusammengestellt.

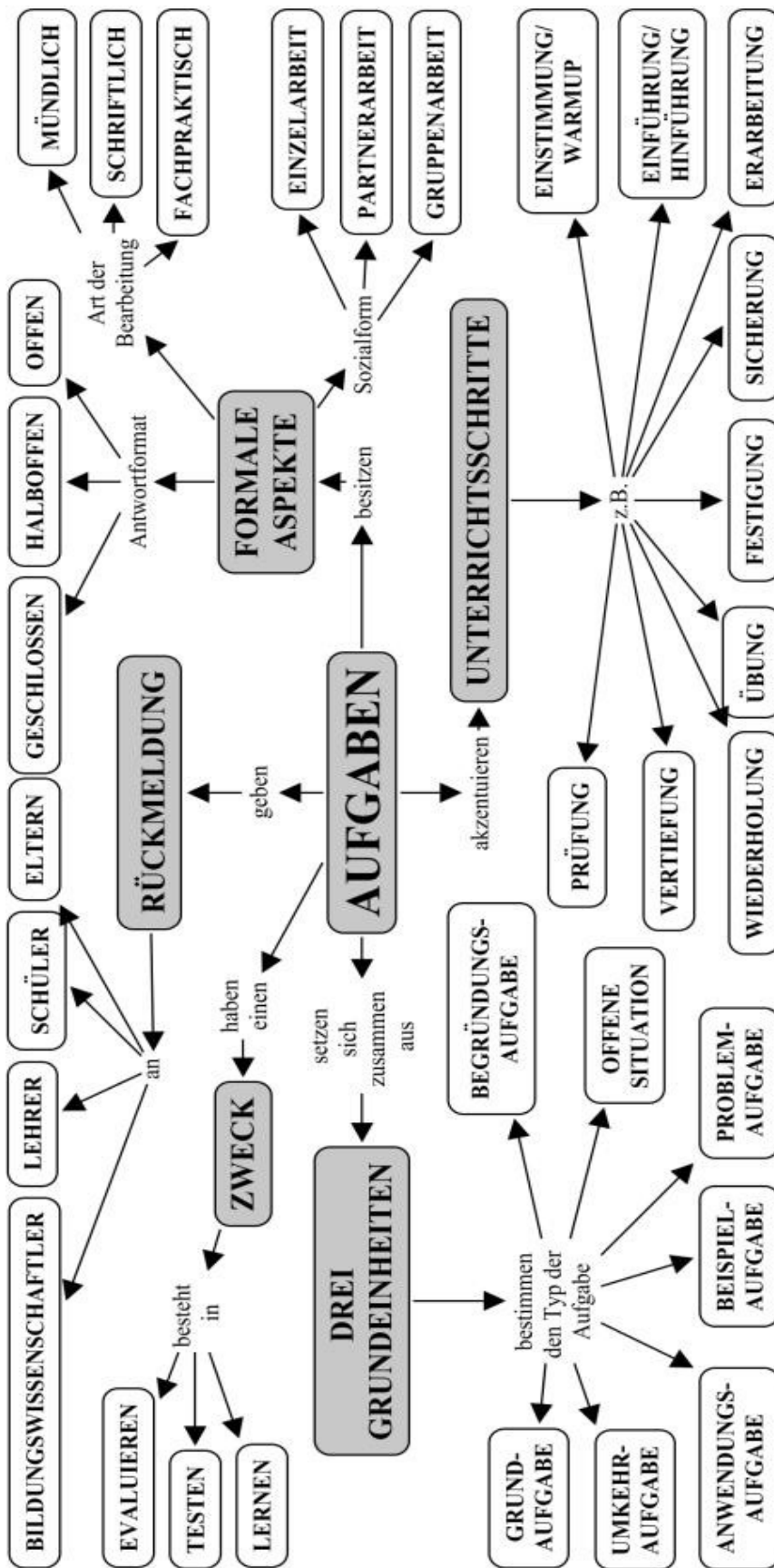


Abb. 14: Zusammenstellung wesentlicher Aspekte von Aufgaben im Unterricht.

3.5.4 Qualitätskriterien für Lernaufgaben

Etwa zwei Drittel der Zeit im Biologieunterricht wird mit Lernaufgaben verbracht, daher sollte diese Zeit möglichst effektiv genutzt werden (Jatzwauk et al. 2008, S. 279). Einige große Forschungsprojekte hatten das Ziel, die Aufgabenkultur im Biologieunterricht zu verbessern (zum Beispiel SINUS (Blume 2000) oder BiK (Bayehuber 2007)) und machten dabei vor allem Probleme in der Schulpraxis deutlich. Ausgehend von diesen Problemen wird im Folgenden auf die Qualität von Aufgaben eingegangen und Kriterien zur didaktische Analyse von Aufgaben aufgeführt.

Aufgabenkultur

Bei einer SINUS-Bedarfsanalyse zur Verbesserung von Lernaufgaben im Unterricht wurden gravierende Mängel bei dem Aufbau von Aufgaben, bei deren Einsatz im Unterricht und bei dem Umgang der Schüler mit diesen Aufgaben festgestellt (Blume und Rademann 2000). Die Probleme sind im Folgenden kurz aufgeführt:

- **Probleme bei der Aufgabenkonstruktion:** Viele Aufgaben haben keinen Bezug zur Lebenswelt und das Vorwissen der Schüler wird kaum berücksichtigt. Die Aufgabenstellung ist standardisiert und oft unklar formuliert. Der Schwierigkeitsgrad ist meist gleichbleibend und es gibt keine differenzierenden Angebote für leistungstärkere und leistungsschwächere Schüler. Die eingesetzten Aufgaben sollen oftmals nur in Einzelarbeit gelöst werden und sind häufig nicht miteinander vernetzt. Meistens gilt nur eine Lösung als richtig.
- **Probleme beim Einsatz im Unterricht:** Aufgaben sind häufig auf einfaches Wiederholen von Routinen angelegt. Sie prüfen nur kurzfristig Gelerntes und zwar meistens genau das, was im vorangegangenen Unterricht erarbeitet wurde. Werden Aufgaben tatsächlich im Unterricht eingesetzt, sind sie oft reine Wiederholungsaufgaben, ohne Verknüpfung des Bekannten mit Neuem. Die Lernenden dürfen im Unterricht häufig zu wenig selbst erarbeiten. Ferner dominiert der fragend-entwickelnde Unterricht, welcher aber oft zu kleinschrittigen Frage-Antwort-Gesprächen zwischen der Lehrkraft und einzelnen Schülern gerät. Des Weiteren wird den Schülern oft zu wenig Zeit für die Lösung der Aufgabe gegeben. Die anspruchsvolleren Aufgaben werden kaum im Unterricht bearbeitet, sondern treten zum überwiegenden Teil in den Abschlusstests auf.
- **Probleme von Schülern beim Umgang mit Aufgaben:** Die Schüler erfahren selten einen Kompetenzzuwachs. Ferner können sie kaum selbständig arbeiten und recherchieren. Die Heterogenität zwischen den Schülern ist sehr groß und es mangelt ihnen erheblich an mündlicher und schriftlicher Ausdrucksfähigkeit.

Nach Thonhauser sind die im Unterricht eingesetzten Aufgaben darüber hinaus meist unzureichend in Kontexte eingebettet und es besteht meistens nur eine geringe Anzahl an Aufgaben, die das selbe Lehrziel betreffen, wodurch es kaum zur Wiederholung und damit zur Festigung des Gelernten kommt. Ferner zielen Aufgaben selten auf vernetztes Wissen ab und gehen häufig mit unklaren Erwartungen bezüglich ihrer Antworten einher. Abschließend bilden diese Aufgaben meist nur den inhaltlichen Aspekt der Lehrziele ab und bedienen sich bei der Auswertung meist unzureichender Bewertungsschlüssel (Thonhauser 2008, S. 22). Das führt dazu, dass sich Schüler mit traditionellen, meist monotonen und unverbundenen Aufgabensammlungen beschäftigen, was letztendlich zu einer schlichten Automatisierung grundlegender Routineverfahren führt und nicht zu einem besseren Verstehen der Sache (Kühn 2011, S. 36).

Allgemeine Aspekte zur Aufgabengüte

Um den aufgeführten Mängeln bei der Konstruktion von Aufgaben zu entgehen, müssen für den reflektierten Umgang mit Aufgaben Analysekriterien erstellt werden, die allgemeingültige Aussagen über die Qualität von Aufgaben zulassen.

Hierbei gilt, dass der Lernzuwachs stark von der Qualität der Aufgaben abhängt (Thonhauser 2008, S. 18). Der Lernzuwachs ist umso höher, desto mehr auf systematisch und theoretisch begründete Aufgabenentwicklung geachtet wird und Lernaufgaben nicht rein intuitiv konstruiert werden (Narciss und Proske 2001, S. 3). Als generelle Merkmale guter Aufgaben kann gesagt werden, dass sie die Eigenaktivität der Lernenden erhöhen und durch das Herstellen von Sinn- und Bedeutungsbezügen deren Motivation steigern (Müller und Helmke 2008, S. 41). Des Weiteren ist eine gute Aufgabe eine Herausforderung an die Lerntiefe der Schüler und fördert gleichzeitig sowohl deren inhalts- als auch prozessbezogene Kompetenzen (Leisen 2005, S. 306). Ferner zeichnet sie sich durch kognitive Aktivierung aus, ist authentisch und lässt Differenzierungen zu (Leuders 2009, S. 7). Ausgehend von Blömeke werden abschließend in Tab. 10 sowohl fachliche als auch didaktische Merkmale von guten Aufgaben aufgeführt und ergänzt.

Tab. 10: Merkmale didaktischer und fachlicher Aufgabenqualität mit Analysekriterien als Erweiterung zu Blömeke (Blömeke et al. 2006, S. 337).

Fachliche und didaktische Merkmale für hohe Aufgabenqualität	Analysekriterien
Exemplarische Erschließung eines relevanten Bildungsinhaltes	Wird in der Aufgabe eine relevante Grundfrage oder Methode thematisiert?
Fachliche Richtigkeit ¹³	Sind die dargestellten Inhalte in der Aufgabe und den Materialien fachlich korrekt?
Ansprache eines Bedürfnisses der Schüler	Wird durch die Aufgabe ein schülerrelevantes Thema angesprochen oder Inhalte aus der Lebenswelt der Schüler erarbeitet?
Textverständlichkeit ¹³	Sind die Fach- und Fremdwörter bekannt bzw. ausreichend erklärt? Sind diese sinnvoll oder erschweren sie den Zugang?
Förderung genereller intellektueller Fähigkeiten	Beansprucht die Aufgabe geforderte kognitive Prozesse oder Wissensformen?
Eigeninitiative und kognitive Aktivierung ⁹	Müssen die Schüler über eingeübte Routinen hinaus (z. B. Informationen aus Text entnehmen) Leistungen erbringen? Welche?
Neuigkeitswert auf Wissens- und Erfahrungsstand (z. B. Basiskonzepte) ¹¹	Ist der Inhalt oder die Methode wirklich neu? Muss zur Lösung Basiswissen (neu) kombiniert werden?
Chance auf Bewältigung	Ist der Grad des spezifischen Inhaltes, der Methode oder der sprachlichen Komplexität der Aufgabenstellung machbar?
Herausforderung / Lertiefe ¹⁰	Ist die Aufgabe über den Grad der Bewältigung hinaus zu trivial oder zu anspruchsvoll?
Differenzierungspotenzial	Ist bei der Aufgabe die Bearbeitung auf unterschiedlichen kognitiven Niveaus, in unterschiedlichen Tiefen oder in unterschiedlichem Umfang möglich?
Authentische Situation	Ist die Aufgabe glaubwürdig?
Herstellen von Bedeutungszusammenhängen ⁹	Sind die Aufgaben mit- und untereinander in irgendeiner Weise verknüpft?
Problemlösefähigkeit fördern	Müssen die Schüler den Lösungsweg selbst finden oder ist die Aufgabe stark vorstrukturiert? Muss zur Lösung fächerübergreifendes Wissen genutzt werden? ¹¹
Erfordernis sozialer Interaktion	Erfordert die Aufgabe wirklich Gruppen- oder Partnerarbeit? Besteht durch die Aufgabe Bedarf an Diskussion oder Reflexion im Klassenverband?
Klare Aufgabenstellung ¹¹	Wird den Schülern klar, was von ihnen erwartet wird?
Einsatz von Fachsprache ¹¹	Wird die Nutzung von Fachsprache gefördert?
Direktes Feedback	Ermöglicht es die Aufgabe, einem Schüler direkte Rückmeldung zu seiner erbrachten Leistung zu geben?
Rahmenbedingungen ¹²	Ist den Schülern bekannt, wieviel Zeit ihnen für die Aufgabe zur Verfügung steht, ob sie dafür benotet werden oder welches Material sie benutzen dürfen und wo sie dieses finden?
Organisation ¹³	Wissen die Schüler, wie sie ihre Ergebnisse am Ende darstellen sollen? Wissen sie ihre Funktion bei Gruppenarbeit?

⁹ Müller und Helmke 2008

¹⁰ Leisen 2005

¹¹ Rademann und Blume 2001

¹² Thonhauser 2008

¹³ Weitzel 2012

Biologiedidaktische Aspekte der Aufgabenqualität

Abgesehen von den angesprochenen Merkmalen und Analyse Kriterien sollten biologische Fachinhalte immer in einen Kontext eingebettet werden. Das ist nachgewiesenermaßen motivierend für Jugendliche und beeinflusst deren Interesse an den Aufgaben (Hammann 2006, S. 94, Weitzel 2012). Gerade in den Naturwissenschaften und vor allem in der Biologie lassen sich Inhalte durch Aufgaben leicht in relevante Zusammenhänge setzen. Entsprechende Möglichkeiten für Kontexte werden in Tab. 11 aufgeführt.

Tab. 11: Interessengenerierende Kontexte in Anlehnung an Hammann (Hammann 2006).

Nimmt die Aufgabe Bezug auf Themen des alltäglichen Lebens (z. B. Alltag, Zuhause, Schule, Sport, Einkaufen, Gesundheit, Krankheit, Ernährung)?
Bezieht sich der Kontext auf die Erde oder Umwelt (z. B. natürliche Ressourcen, Erhalt und Gefährdung unserer Umwelt, Boden, Wetter, Klima)?
Besteht eine Verwendung in der Technologie (z. B. Biotechnik, Gentechnik, Energieverwertung, Transport, Abfall, Materialverwertung)?
Entspricht der Kontext der Wissensgenese (z. B. historische und aktuelle Forschungen)?

Nach Hammann lassen sich alle vier thematischen Bezüge in einen bedeutsamen personalen, sozialen oder globalen Zusammenhang setzen. So sollte Schülern aus einer Aufgabe klar werden, welche Konsequenzen das Thema für sie persönlich, ihre Familie, die Gesellschaft oder die Erde hat.

Aufgabenqualität und Einsatz von Aufgaben im Unterricht

Abgesehen von den bisherigen Qualitätsmerkmalen, die sich direkt auf Aufgaben beziehen, gibt es einen weiteren Punkt, der ausschlaggebend für den Erfolg einer Aufgabe ist: der Einsatz der Aufgabe im Unterricht. So kommt es im schulischen Kontext für den Lernerfolg nicht nur darauf an, möglichst gute Aufgaben zu konstruieren, sondern auch auf die Art, wie man mit diesen Aufgaben im Unterricht umgeht (Bruder 2003, S. 12). Gute Aufgaben sind demnach nicht von selbst gute Aufgaben, sondern nur so gut wie ihre Tauglichkeit in der jeweiligen Lehrsituation (Leuders 2009, S. 8). Aus diesem Grund ist es immer ratsam, sich bei der Bewertung nicht nur generelle Qualitätskriterien von Aufgaben klar zu machen, sondern auch, in welcher Lernphase sich die Aufgabe im Unterricht befindet (Leuders 2009, S. 11). Ferner muss der Einsatz guter Aufgaben im Unterricht didaktisch arrangiert werden und Ein- beziehungsweise Ausstiege oder Überleitungen zu weiteren Aufgaben erstellt werden (Leisen 2006, S. 266; Winter 2008, S. 118). Da dies sehr stark in die Unterrichtsplanung führt und nicht Thema dieser Arbeit ist, werden im Folgenden nur allgemeine Kriterien für den gelungenen Einsatz im Unterricht aufgeführt.

Generell sollte die Platzierung der Aufgabe die Selbstständigkeit der Schüler fördern. Hierfür müssen notwendige Basisinformationen vorhanden sein und ausreichend Zeit für die schülerverantwortliche Bearbeitung eingeplant werden. Ferner sollte eventuell eine schriftliche Lösungshilfe beziehungsweise eine aufbauende, variierende Folge- oder Hausaufgabe vorbereitet sein (Rademann und Blume 2001, S. 432 f.).

Formale Aspekte einer klaren Aufgabenstellung

Ein weiterer sehr wichtiger Faktor für die Qualität von Aufgaben ist eine klare Aufgabenstellung. Aus der Aufgabenstellung sollte hervorgehen, was genau gemacht werden soll. Für das Fach Biologie wurden Operatoren formuliert, definiert und in der Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA) festgehalten (KMK 2004). Operatoren dienen als Signalwörter (zum Beispiel: *zeichne* oder *vergleiche*), die den Lernenden vorgeben, was zur Erfüllung der Aufgabe von ihnen erwartet wird. Damit Lernende diese Hinweise verstehen und angemessen umsetzen, müssen diese im Vorfeld angewandt und geübt werden. Das erfordert jedoch die einheitliche Verwendung dieser Operatoren und deren Umsetzung. Dass dies nicht ganz einfach ist, soll das folgende Beispiel verdeutlichen. Was wird mit den folgenden Operatoren in einer Antwort ganz genau gefordert?

Begründe / Erkläre / Erläutere / Diskutiere den Sachverhalt!

Um Klarheit zu schaffen sind im Folgenden die Operatoren mit den dazugehörigen erwarteten Leistungen aufgeführt (KMK 2004):

- **Begründe:** Schüler sollen Sachverhalte auf Regeln und Gesetzmäßigkeiten bzw. kausale Beziehungen von Ursachen und Wirkungen zurückführen.
- **Erkläre:** Schüler sollen einen Sachverhalt auf Regeln und Gesetzmäßigkeiten zurückführen sowie ihn nachvollziehbar und verständlich darstellen.
- **Erläutere:** Schüler sollen einen Sachverhalt veranschaulichend darstellen und durch zusätzliche Informationen verständlich machen.
- **Diskutiere:** Schüler sollen Argumente und Beispiele zu einer Aussage oder These einander gegenüberstellen und abwägen.

Es wird deutlich, dass Verständnisprobleme und unklare Anforderungserwartungen bei der Aufgabenstellung zu unerwünschten Ergebnissen führen, die nichts mit dem Wissen und Können der Lernenden zu tun haben, sondern mit der schlichten Wortwahl in der Aufgabenstellung. Eine verständliche Arbeitsanweisung und damit verbundene

Zieltransparenz erleichtert den Lernenden das Lösen einer Aufgabe erheblich (Helmke 2006b).

Um eine Aufgabe verständlicher zu machen, kann diese stark vorstrukturiert werden. Die Entscheidung ist von vielen Faktoren abhängig, wie zum Beispiel dem konkreten Lehrziel, dem Alter und der Reife der Schüler oder der Situation in der unterrichtet wird. Nichtsdestotrotz brauchen Lernende ein Mindestmaß an fachlichen, inhaltlichen und unterrichtsmethodische Strukturen an denen sie sich orientieren oder auf die sie im Zweifelsfall zurückgreifen können (Leisen 2006, S. 264). Ganz praktische Handlungsanweisungen für den gelungenen Einsatz von Aufgaben gibt Prediger. Den Schülern sollte vor der Aufgabe klargemacht werden: „Was kann ich in dieser Einheit lernen? Was muss ich erstellen/zeigen? Was muss ich tun, um dorthin zu gelangen? Wo finde ich, was ich zum Arbeiten brauche?“ (Prediger 2008, S. 6).

Nach der Bearbeitung der Aufgabe sollten mehrere Lernende zusammenhängend ihre Lösungsvorschläge vorstellen und dabei die Fachsprache richtig einsetzen (Rademann und Blume 2001, S. 432 f.). Hierbei sollen Fehler zwar zugelassen, aber zeitnah thematisiert und sachorientierte Rückmeldung gegeben werden (Jacobs 2008, S. 110; Astleitner 2008, S. 69).

3.5.5 Schwierigkeitsgrade von Aufgaben

Den Schwierigkeitsgrad einer Aufgabe einzuschätzen oder bewusst zu generieren, ist nicht einfach. Das tägliche Geschäft der Erstellung von Test- und Lernaufgaben fällt dabei nicht nur Lehrenden im Schuldienst schwer (McElvany et al. 2009, S. 232, Luthiger 2008, S. 38), sondern auch Fachdidaktikern. Dies liegt nicht nur an den vielen Ebenen von Aufgaben im Unterricht, sondern auch an der Vielfalt an Möglichkeiten den Schwierigkeitsgrad einer Aufgabe zu beschreiben. In den folgenden Abschnitten wird eine Auswahl an Möglichkeiten vorgestellt, die in der Diskussion von Aufgabenschwierigkeit momentan verwendet werden. Ein Vorgriff auf die hier vorgestellten Methoden zeigt Abb. 15.

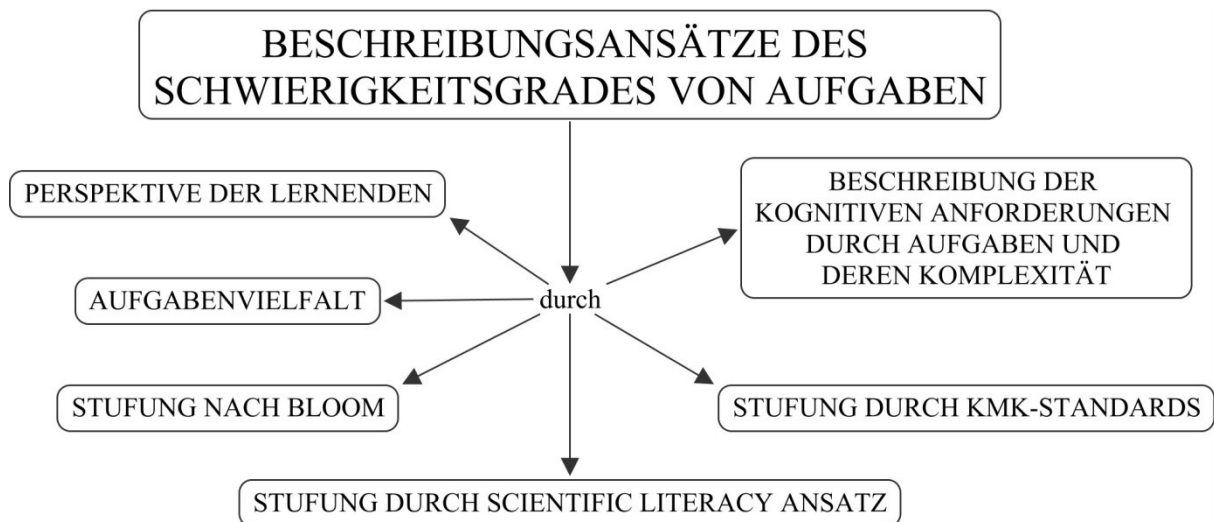


Abb. 15: Beschreibungsansätze für den Schwierigkeitsgrad von Aufgaben.

Aufgabenschwierigkeit aus Lerner-Perspektive

Aus Sicht der Lernenden lässt sich nach Nunan (Nunan und Keobke 1995, S. 2 f.) die Schwierigkeit einer zu lösenden Aufgabe in drei Kategorien aufteilen: Perspektive der Lernenden, Text und Aufgabe. Für Lernende wird die Aufgabe unter anderem leichter, wenn sie zuversichtlich sind, die Aufgabe zu bewältigen, über Vorwissen und die nötigen sprachlichen Fähigkeiten verfügen und ausreichend Zeit haben. Der Text einer Aufgabe macht es Lernenden leichter, wenn er wenig detaillierte Fakten beinhaltet, klar strukturiert ist, vertraute Inhalte behandelt, Informationen sequenziert präsentiert, graphische und non-verbale Hilfen enthält und keine zeitlichen Sprünge macht. Ferner hat sich gezeigt, dass die Verwendung von Fachbegriffen einen Einfluss auf die Aufgabenschwierigkeit hat (Schmiemann 2012, S. 129). Abgesehen von den Lernenden und Texten gibt es nach Nunan Anhaltspunkte für den Schwierigkeitsgrad bei den Aufgaben selbst. Eine Aufgabe fällt Lernenden leichter, wenn sie von geringer kognitiver Komplexität ist, aus wenigen Schritten besteht, in Kontexte eingebunden ist, Lernhilfen verfügbar sind und ausreichend Bearbeitungszeit zu Verfügung steht.

Aufgabenschwierigkeit durch Aufgabenvielfalt

Nach Mayer (Mayer 2001, S. 52) kann die Einteilung von Aufgaben in Schwierigkeitsgrade aber auch anhand von Schülerleistung erfolgen. Das geschieht entsprechend der Frage, was Lernende zum erfolgreichen Lösen der Aufgabe leisten müssen. Hierzu gehört zum Beispiel das Erklären, Vergleichen, Beschreiben, Aufzählen und Ordnen von Inhalten. Im Gegensatz zu Bloom (siehe folgenden Absatz) beinhalten diese Operationen keine Stufung, sondern beleuchten eine andere Facette der Aufgaben, nämlich deren Vielfalt und Anzahl. Nach Mayer bedeutet das, je unterschiedlicher die beanspruchten Operationen durch Aufgaben zum gleichen Thema sind, desto besser. Die Schwierigkeit der Aufgaben lässt sich hierbei nach Mayer nur bedingt einschätzen, da unterschiedliche Lernende auch unterschiedliche Stärken und Schwächen haben. Durch die bewusste Variation der Aufgaben kommen jedoch alle Schüler irgendwann an den Punkt an dem eine Aufgabe für sie schwieriger wird und sie mehr leisten müssen.

Schwierigkeitsgrade nach Bloom

Im Gegensatz zu Nunans und Mayers Definition der Aufgabenschwierigkeit, lassen sich Aufgaben auch aus lernpsychologischer Sicht klassifizieren. Die international bedeutendste und am gründlichsten diskutierte Klassifizierung des Schwierigkeitsgrades stammt von Benjamin Bloom (Germ und Harms 2010, S. 5). Bei der Beforschung des Lernens unterschied Bloom drei relevante Gebiete: den kognitiven, den affektiven und den psychomotorischen Bereich. Im Folgenden wird auf die Lernzieltaxonomie des kognitiven Bereiches eingegangen, der sich folgendermaßen zusammensetzt (Bloom und Engelhart 1976):

- 1) Kenntnisse und Wissen: Lernende kennen Einzelheiten wie Begriffe, Definitionen, Fakten, Daten, Regeln, Gesetzmäßigkeiten, Theorien, Merkmale, Kriterien, Abläufe. Lernende können Wissen darüber abrufen und wiedergeben.
- 2) Verstehen: Lernende können Sachverhalte mit eigenen Worten erklären oder zusammenfassen. Sie können Beispiele anführen, Zusammenhänge verstehen und Aufgabenstellungen interpretieren.
- 3) Anwenden: Lernende können das Gelernte und ihr Wissen problemlösend einsetzen und in neuen Situationen anwenden. Ferner können sie unaufgefordert abstrahieren oder Abstraktionen verwenden.
- 4) Analyse: Lernende können ein Problem in einzelne Teile zerlegen und somit die Struktur des Problems verstehen. Des Weiteren können sie Widersprüche aufdecken, Zusammenhänge erkennen und Folgerungen ableiten. Außerdem können sie zwischen Fakten und Interpretationen unterscheiden.

- 5) Synthese: Lernende können aus mehreren Elementen eine neue Struktur aufbauen oder eine neue Bedeutung erschaffen. Sie können neue Lösungswege vorschlagen, neue Schemata entwerfen oder Hypothesen stellen und diese begründen.
- 6) Beurteilung: Lernende können den Wert von Ideen und Materialien beurteilen und Alternativen gegeneinander abwägen. Sie können auswählen, Entschlüsse fassen und diese begründen. Sie können dabei bewusst Wissen transferieren, zum Beispiel durch Arbeitspläne.

Schwierigkeitsgrade nach dem Scientific Literacy Ansatz

Einen spezielleren und eher fachlichen Rahmen für die Graduierung von Aufgaben bietet der Scientific Literacy Ansatz (Bybee 1997). Hierbei wird nicht über explizites Wissen, sondern über Kompetenz eine Definition von naturwissenschaftlicher Grundbildung gegeben. Die Kompetenz beschreibt dabei die fachunabhängige Fähigkeit, naturwissenschaftliche Fragestellungen zu erkennen und Wissen anzuwenden, Phänomene zu beschreiben und Schlussfolgerungen zu ziehen sowie Bereitschaft zu zeigen sich mit naturwissenschaftlichen Themen zu beschäftigen und lebenslang damit auseinander zu setzen. Die naturwissenschaftliche Kompetenz bezieht sich hierbei nicht nur auf Wissen und Interesse an Naturwissenschaften, sondern auch auf Wertschätzung und Verantwortung gegenüber anderen, der Umwelt und der natürlichen Ressourcen. Speziell in Bezug auf Aufgaben lassen sich nach Schecker hieraus vier Stufen einer naturwissenschaftlichen Bildung in Bezug auf Bybee formulieren (Schecker und Parchmann 2006, S. 49):

- 1) Die nominale Scientific Literacy (SL): die Kenntnis über naturwissenschaftliche Themenstellungen und Begriffe, die jedoch im wissenschaftlichen Sinne falsch verstanden werden.
- 2) Die funktionale SL: Faktenwissen und konkrete Verwendung von naturwissenschaftlichem Vokabular und von Formalismen.
- 3) Die konzeptionelle und prozedurale SL: das Verständnis zentraler naturwissenschaftlicher Ideen und Verfahren und der Herstellung von Beziehungen zwischen Fakten, Begriffen und Prinzipien.
- 4) Die multidimensionale SL: Das Verständnis der Besonderheiten naturwissenschaftlichen Denkens und die Fähigkeit zur Einordnung in soziale und kulturelle Zusammenhänge.

Diese vier Punkte beschreiben eine mögliche Stufung des Schwierigkeitsgrades, allerdings sind sie wenig konkret und lassen für eine konkrete Klassifikation von Aufgaben zu viel Freiheiten und Spielraum.

Schwierigkeitsgrade nach den KMK-Standards

Die Klassifikation von Aufgaben durch Kompetenzen steht seit einiger Zeit im Fokus der biologiepädagogischen Forschung. Hierbei werden Aufgaben über die Kompetenzen beschrieben, die zur Lösung der Aufgabe notwendig sind (Hammann 2006, S. 85). Die grobe Bestimmung der Aufgabenschwierigkeit erfolgt hierbei auch nicht nach Oberflächenmerkmalen wie bei den ersten beiden Ansätzen, sondern über das zweidimensionale Kompetenzstrukturmodell der allgemeinen Bildungsstandards (2004) in Tab. 12 auf Seite 81 (Schaal und Weitzel 2012, S. 141). In Anlehnung an die einheitlichen Prüfungsanforderungen der Abiturprüfung (KMK 2004) befinden sich auf der X-Achse die Kompetenzbereiche der Biologie und auf der Y-Achse deren Anforderungsbereiche.

Die Anforderungsbereiche haben explizit keinen Stufencharakter und erlauben ausdrücklich keine Graduierung von Kompetenzen (Kauertz et al. 2010). Bei den Anforderungsbereichen handelt es sich ausschließlich um ein normativ festgelegtes Maß kognitiver Anforderungen, das Lernende zur Lösung einer Aufgabe erbringen müssen (Schaal und Weitzel 2012). Diese kognitiven Anforderungen werden weitläufig verstanden als Reproduktion, Anwendung und Transfer. Reproduktion bedeutet, dass keine neuen Informationen erzeugt werden müssen, also ein unveränderter Abruf von Wissen, Methoden und Fertigkeiten aus dem Langzeitgedächtnis. Anwendung liegt dann vor, wenn die Aufgabe Hinweise enthält, wie das bereits Erlernte eingesetzt werden soll oder wenn aus bereits vorhandener Information ein logischer Schluss gezogen werden muss. Das Überführen von vorhandenen Informationen oder Kompetenzen in eine neue Situation oder Kontext hingegen kennzeichnet den Transfer.

Tab. 12: Kompetenzmatrix der Bildungsstandards nach Leisen (Leisen 2006, S.261).

		Kompetenzbereiche			
		Fachwissen	Erkenntnisgewinnung / Fachmethoden	Kommunikation	Bewertung / Reflexion
Anforderungsbereiche	I	Kenntnisse und Konzepte zielgerichtet wiedergeben	Bekannte Untersuchungsmethoden und Modelle beschreiben, Untersuchungen nach Anleitung durchführen	Bekannte Informationen in verschiedenen fachlich relevanten Darstellungsformen erfassen und wiedergeben	Vorgegebene Argumente zur Bewertung eines Sachverhaltes erkennen und wiedergeben
	II	Kenntnisse und Konzepte auswählen und anwenden	Geeignete Untersuchungsmethoden und Modelle zur Bearbeitung überschaubarer Sachverhalte auswählen und anwenden	Informationen erfassen und in geeigneten Darstellungsformen situations- und adressatengerecht veranschaulichen	Geeignete Argumente zur Bewertung eines Sachverhaltes auswählen und nutzen
	III	Komplexere Fragestellungen auf der Grundlage von Kenntnissen und Konzepten planmäßig und konstruktiv bearbeiten	Geeignete Untersuchungsmethoden und Modelle zur Bearbeitung unbekannter Sachverhalte begründet auswählen und anpassen	Informationen auswerten, reflektieren und für eigene Argumentationen nutzen	Argumente zur Bewertung eines Sachverhaltes aus verschiedenen Perspektiven abwägen und Entscheidungsprozesse reflektieren

Schwierigkeitsgrade nach Komplexität und kognitiven Prozesse

Für die Erstellung von gezielt anspruchsvollen Aufgaben werden nach Kauertz schwierigkeitszeugende Merkmale wie Komplexität und kognitive Prozesse in einem Modell vereint und mit den entsprechenden Kompetenzbereichen verbunden (Wellnitz und Mayer 2008, S. 133, Kauertz et al. 2010, S. 139). Die Leitidee stellen die Bildungsstandards, die sich für die Naturwissenschaften in einigen Punkten überschneiden, aber dennoch disziplinspezifisch sind. Sie werden durch kognitive Aktivitäten ergänzt, die zur erfolgreichen Lösung der Aufgabe genutzt werden. Der Anstieg in der Komplexität geht in der Regel mit einer Zunahme von Lösungsschritten einher, die von der Nennung einzelner Elemente bis hin zur Verknüpfung dieser Elemente reicht (Wellnitz et al. 2012). Zur Beschreibung der Leistung innerhalb der Komplexität wird eine Skala herangezogen. Auf dieser Skala werden sechs hierarchisch geordnete Stufen beschrieben: (I) ein Fakt, (II) mehrere Fakten, (III) ein Zusammenhang, (IV) unverbundene Zusammenhänge, (V) mehrere verbundene Zusammenhänge oder (VI) ein übergeordnetes Konzept (Neumann et al. 2007, S. 111 f.). Das daraus resultierende dreidimensionale Kompetenzmodell bildet die Grundlage für die Vorhersage der

Schwierigkeit von Aufgaben und deren gezielter Konstruktion (siehe Abb. 16). Diese erfolgt, indem die Dimensionen der Komplexität, der kognitiven Prozesse und der Bildungsstandards untereinander in Beziehung gesetzt werden (Leutner et al. 2008, Wellnitz et al. 2012, S. 269).

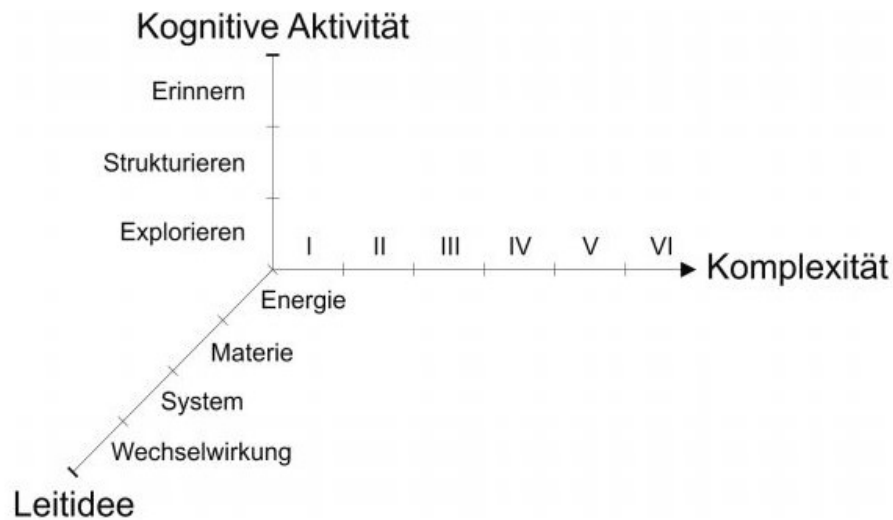


Abb. 16: Vorschlag eines Strukturmodells physikalischer Kompetenz als Basis zur gezielten Generierung des Schwierigkeitsgrades bei Aufgaben (entnommen aus Neumann (2007), Seite 111).

Die Überprüfung dieses Modells im Bereich der Physik hat gezeigt, dass die Komplexitätsstufen und die kognitiven Prozesse mittlere bis starke Effekte auf die Aufgabenschwierigkeiten haben (Wellnitz et al. 2012, S. 282). Allerdings hing die Schwierigkeit der eingesetzten Aufgaben nicht nachweisbar mit den Leitideen und dadurch mit dem Bildungsplan zusammen (Neumann et al. 2007, S. 113). Die allgemeine Formulierung dieses Modells erlaubt dessen Übertragung auf die einzelnen naturwissenschaftlichen Fächer. Die Überprüfung des Kompetenzstufenmodells hat gezeigt, dass sich das Modell für die Prognose des Schwierigkeitsgrades von Aufgaben in den naturwissenschaftlichen Fächern eignet (Wellnitz et al. 2012, S. 282). Um also gezielt Aufgaben mit steigender Schwierigkeit und Komplexität für den Einsatz im Biologieunterricht zu konstruieren, sollte das Kompetenzstufenmodell auf das Fach übertragen werden. Die biologiespezifische Anwendung der Kompetenzstufen nach Mayer (2013) wird in Tab. 13 aufgeführt.

Tab. 13: Kompetenzstufen im Bereich Fachwissen nach Mayer (Mayer et al. 2013).

Kompetenzstufe I: Lernende können einzelne biologische Fakten und Fachbegriffe beziehungsweise einfache biologische Sachverhalte innerhalb einer lebensweltlichen Situation oder innerhalb eines übersichtlichen Fachtextes lokalisieren und wiedergeben.

Kompetenzstufe II: Über die bereits auf Stufe I vorhandenen Kompetenzen hinaus können Lernende auf Kompetenzstufe II (Mindeststandard) einfache biologische Zusammenhänge mit eigenen Worten darstellen und in einfachen Kontexten anwenden.

Kompetenzstufe III: Auf Kompetenzstufe III (Regel-Standard) können Lernende biologische Zusammenhänge beschreiben und erklären sowie biologische Konzepte und Prinzipien anwenden.

Kompetenzstufe IV: Lernende der Kompetenzstufe IV (Regelstandard plus) können über die vorhergehenden Kompetenzstufen hinaus komplexe biologische Zusammenhänge unter Anwendung von Konzepten und Prinzipien erklären.

Kompetenzstufe V: Lernende auf der Kompetenzstufe V (Optimal-Standard) bewältigen nicht nur mit großer Sicherheit die Anforderungen der Kompetenzstufen I bis IV, sondern können darüber hinaus ihnen zumeist unbekannte komplexe biologische Zusammenhänge unter selbstständiger Anwendung von Konzepten und Prinzipien erklären und begründen.

Aus Komplexität erwächst Schwierigkeit

Der Schwierigkeitsgrad einer Aufgabe lässt sich über die kognitiven Anforderungen hinaus über Inhalt(e) bestimmen. Je breiter eine Aufgabe inhaltlich aufgestellt ist, desto komplexer wird sie. Die Erhöhung der Anzahl der Inhalte geht dabei aber nicht zwangsläufig mit einer Steigerung des Schwierigkeitsgrades einher, sondern kann in der schlichten Verlängerung der Bearbeitungszeit münden (Aufschnaiter 2003, S. 5). Ferner spielt das Vorwissen der Schüler eine Rolle. Je weniger Schüler zu einem Thema wissen, desto schwieriger werden sie die entsprechende Aufgabe empfinden (Schaal und Weitzel 2012, S. 140).

Die unbestritten einfachste Art und Weise, den Schwierigkeitsgrad einer Aufgabe zu ermitteln, ist diese durch möglichst viele Schüler bearbeiten zu lassen und im Nachhinein zu überprüfen, wie oft sie richtig gelöst wurde. Aus den Antworten der Schüler lässt sich daraufhin die Lösungswahrscheinlichkeit einer Aufgabe ableiten und daraus deren Schwierigkeitsgrad. Je geringer die Lösungswahrscheinlichkeit, desto schwieriger die Aufgabe (Astleitner 2008, S. 65). Dieses Vorgehen ist plausibel, für die gezielte Konstruktion von unterschiedlich schweren und leistungsdifferenzierenden Abituraufgaben jedoch unbefriedigend.

3.5.6 Überblick der Perspektiven auf den Schwierigkeitsgrad

Insgesamt werden für diese Arbeit folgende Ansätze zur Beschreibung des Schwierigkeitsgrades von Aufgaben berücksichtigt:

- 1) Bloom (1976) setzt bei der Einschätzung von Aufgaben das Wissen in den Vordergrund, indem er den Umgang mit Wissen als eine Stufung beschreibt. Die Stufung geht über das pure Wissen zum Verstehen, zur Anwendung, zur Analyse, zur Synthese bis hin zu einer begründeten Beurteilung.
- 2) Ein anderer, allgemein gültiger Ansatz wird von Mayer (2001) beschrieben. Mayer geht bei seiner Einschätzung des Schwierigkeitsgrades nicht von einem Schüler aus, sondern von der Unterschiedlichkeit von mehreren Lernenden in Bezug auf deren Fähigkeiten. Er beschreibt einen lernwirksamen Schwierigkeitsgrad nicht als gezielt einschätzbare Möglichkeit, sondern als Annäherung an das jeweilige Schülerniveau durch die Steigerung der Aufgabenanzahl und vor allem durch deren Vielfalt. Mayer sagt, dass je unterschiedlicher die Ansprüche durch Aufgaben sind, desto unterschiedlichere Leistungen müssen Schüler erbringen, desto höher ist die Chance, dass ein Schüler dabei an seine persönliche Grenze stößt, was wiederum die Chance auf echten Lernzuwachs erhöht.
- 3) Auch Nunan (1995) geht davon aus, dass alle Schüler unterschiedlich sind. Diese Unterschiedlichkeit in Bezug auf Vorwissen oder sprachliche Fähigkeiten macht die Einschätzung des Schwierigkeitsgrades kompliziert. Je nach persönlichen Voraussetzungen kann ein Schüler zum Beispiel Aufgabe A als schwierig empfinden und Aufgabe B als leicht. Bei einem anderen Schüler könnte die Einschätzung der gleichen Aufgaben umgekehrt ausfallen. Nunan bezieht sich bei der Definition von Schwierigkeit durch Aufgaben auf konkret beschreibbare Merkmale: Aufgabeninhalt, Bearbeitungszeit und Aufgabenmaterial.
- 4) Ähnlich der Taxonomie nach Bloom kann der Scientific Literacy Ansatz nach Bybee (1997) als Stufung von Schwierigkeit beschrieben werden. Im Vergleich zu Bloom bewegt sich dieser allerdings nicht im allgemeinen Bereich, sondern bezieht sich explizit auf die Naturwissenschaften. Ganz konkret werden hier Kompetenzen in der naturwissenschaftlichen Bildung beschrieben. Die Graduierung erfolgt über die nominale Benennung und das Wissen von naturwissenschaftlichen Themen und Begriffen, über deren funktionale Anwendung zu einem konzeptionellen und prozeduralen Verständnis von Fakten und Prinzipien, bis zu einer multidimensionalen Auseinandersetzung von Wissenschaft in sozialen und kulturellen Zusammenhängen.
- 5) Eine Konkretisierung des Schwierigkeitsgrades von Aufgaben im naturwissenschaftlichen Rahmen stellen die Bildungsstandards der KMK (2004) dar. Hierbei

werden je nach Disziplin Kompetenz- und Anforderungsbereiche beschrieben, die die Einordnung einer Aufgabe erleichtern. Ähnlich zu den Bildungsstandards werden neuerdings Aufgaben durch Kompetenz- und Strukturmodelle eingeteilt (Kauertz et al. 2010). Zur Einschätzung von Aufgaben werden hierbei sowohl die Kompetenzbereiche der jeweiligen wissenschaftlichen Disziplinen, als auch die kognitiven Anforderungen an die Schüler, sowie die Komplexität der Aufgaben berücksichtigt. Für die Biologie wurden von Mayer (2013) Kompetenzstufen formuliert, welche konkret die Schülerleistungen auf dem jeweiligen Niveau beschreiben und damit eine Einschätzung des Schwierigkeitsgrades zulassen.

Bei der zusammenfassenden und vergleichenden Darstellung der Facetten von Aufgaben wird deutlich, dass schwierigkeitsgenerierende Merkmale meist aus nur einer Perspektive beschrieben werden (siehe Tab. 14). Hierbei kommt es zu Überschneidungen, was eine Trennung erschwert und bei der konkreten Einteilung von Aufgaben in Schwierigkeitsgraden zu Problemen führt. Aus diesem Grund werden in dieser Arbeit nur zwei Ansätze zur Beschreibung der Schwierigkeit von Aufgaben verwendet: die Bildungsstandards im Fach Biologie der KMK (2004) und das Kompetenzstufenmodell nach Mayer (2013).

Tab. 14: Die dargestellten Kriterien ergeben sich aus den im Text beschriebenen Theorien und sind eine Gegenüberstellung der unterschiedlichen Perspektiven auf den Schwierigkeitsgrad von Aufgaben. Die fett gedruckten Theorien wurden bei der Aufgabenkonstruktion berücksichtigt. Die Klassifikation der Aufgaben dieser Studie befindet sich in Tab. 18 auf Seite 113.

	Bloom	Mayer*	Nunan	Bybee	KMK	Kauertz	Mayer**
Bezieht sich auf Merkmale in Bezug auf den Umgang mit Wissen	ja	ja	ja	ja	ja	ja	ja
Berücksichtigt kognitive Voraussetzungen der Lernenden	-	ja	ja	-	-	-	-
Bezieht sich auf die Art und Merkmale der Aufgabe	-	-	ja	-	-	-	-
Berücksichtigt den Bezug der Aufgabe auf das naturwissenschaftliche Setting	-	-	-	ja	ja	ja	ja
Berücksichtigt die Kompetenzen (Können)	-	-	-	ja	ja	ja	ja

* Richard E. Mayer (2001)

** Jürgen Mayer (2013)

ja = wird in dieser Theorie berücksichtigt | - = wird nicht berücksichtigt

3.6 Differenzierung durch Lernaufgaben

Nach Wodzinski und Wodzinski ist Unterricht besonders dann erfolgreich, wenn er sich an den Bedürfnissen der Lernenden orientiert, an ihre Vorerfahrungen anknüpft, ihr Interesse weckt, im Schwierigkeitsgrad angepasst ist und Raum für individuelle Lernerfahrungen und Lernprozesse ermöglicht (Wodzinski und Wodzinski 2007, S. 4). Für die Anpassung des Unterrichts an die Bedürfnisse der Lernenden sind Lernaufgaben besonders geeignet (siehe Kap. 3.5.3 und Thonhauser 2008, Schaal und Weitzel 2012). Durch unterschiedliche Aufgaben werden darüber hinaus die Lernprozesse der Schüler differenziert angeregt (Bayrhuber 2000, S. 210). Ferner wird das „Könnens-Bewusstsein“ der Schüler gefördert und individuelle Fortschritte nachvollziehbar gemacht (Luthiger 2008, S. 39).

Differenzierung durch Lernaufgaben erfolgt im Unterricht in der Regel auf zwei Wegen. Zum einen ist das die geschlossene Differenzierung, die über die Zuteilung der Aufgaben durch die Lehrkraft erfolgt, und zum anderen durch den selbstdifferenzierenden Ansatz, bei dem sich die Schüler die Aufgaben eigenständig aussuchen und bearbeiten. Auf beiden Wegen sollte der Anspruch jedoch niemals unter einen verbindlichen Mindesthorizont sinken (Prediger 2008, S. 1). Gerade bei der Bearbeitung von anspruchsvolleren Aufgaben ist der Anteil kognitiv aktivierender Unterstützung höher, vor allem wenn Gelegenheiten zur Kooperation geboten werden (Krammer 2009, S. 8). Im Idealfall erfolgt dies bei optimaler Passung zwischen Aufgabenschwierigkeit und individuell vorhandenen Schülerfähigkeiten (Leutner et al. 2008, S. 175, Astleitner 2008, S. 72, Bohl 2012, S. 11).

Bei der planvollen Anpassung von Aufgaben an die Schüler treten allerdings gleich mehrere Probleme in der Schulpraxis auf. Zum einen sind das die teilweise eklatant großen Unterschiede zwischen Fähigkeiten und Vorwissen von Schülern, die nicht statisch sind, sondern sich im Laufe der Zeit ändern können, aber möglichst zu jeder Zeit der Lehrkraft bekannt sein sollten (Bohl 2012). Ferner entsteht durch individualisiertes Lernen und dennoch genormte Prüfungen eine noch größere Leistungskluft zwischen Schülern und damit ein Dilemma für Lehrer, das den meisten Bildungsverwaltungen bislang nicht bewusst zu sein scheint (Winter 2008, S. 121). Neben den positiven Auswirkungen der individualisierten Lernunterstützung muss aber auch das Risiko mitbedacht werden, das Lernen nicht zu sehr zu didaktisieren oder bis zur Oberflächlichkeit zu reduzieren. Die kann nämlich dazu führen, dass es bei Lernenden mit einem hohen Vorwissen zum Thema zu Langeweile und schließlich Desinteresse kommen kann (Seufert und Brünken 2004, S. 144, Kalyuga et al. 2003).

Differenzierung durch Lernaufgaben mit gestuften Hilfen

Der grundsätzliche Sinn von innerer Differenzierung liegt in der Reduktion des Steckenbleibens von Schülern bei der individuellen Bearbeitung von Aufgaben (Wodzinski 2007, S. 18). Dieser Grundsatz liegt unter anderem darin begründet, dass Schüler die Bearbeitung einer Aufgabe nach spätestens 5 Minuten abbrechen, wenn ihnen bis dahin kein erfolgsversprechender Lösungsansatz gelingt (Aufschnaiter 2000, S. 235).

Bei Lernaufgaben mit gestuften Hilfen wird daher eine komplexe Fragestellung (mit möglichst eindeutigem Ergebnis) derart mit Hilfen versehen, dass diese schrittweise zur eigenständigen Erarbeitung der Lösung genutzt werden können (Stäudel 2009a, S. 72). Dabei werden die Anforderungen und die Komplexität des Problems nicht vermindert, sondern Hilfen zur Verfügung gestellt, die ein eigenes Lerntempo ermöglichen und sachbezogene Kommunikation unterstützen (Stäudel 2009a, S. 72). Weitere Vorteile sind, dass Lehrkräfte die Schüler weder zu Aufgaben zuteilen müssen (Gefahr der Diskriminierung), noch dass sich Schüler ihre Aufgaben selber aussuchen müssen (Gefahr der eigenen Über- bzw. Unterschätzung). Hierdurch wird es möglich, den unterschiedlichen Lernvoraussetzungen der Schüler gerecht zu werden und gleichzeitig die Aufgabe an das Fähigkeitsniveau der Schüler anzupassen (Forschergruppe-Universität Kassel et al. 2007, S. 42, Leisen 2006, S. 263).

Für leistungsschwächere Schüler bedeuten Hilfen eine einfache Unterstützungsmaßnahme zur gezielten Beeinflussung des Wissenserwerbs (Brünken et al. 2005, S. 72). Ferner ermöglichen die Hilfen den Komplexitätsgrad von Aufgaben beizubehalten, ohne dass Lernende Gefahr laufen, an der Aufgabe zu scheitern. Das Erbringen von selbstständigen Teilleistungen auf dem Weg zur Lösung verhindert nicht nur das Scheitern an einer Aufgabe. Eigenständige Teilleistungen ermöglichen es auch leistungsschwächeren Schülern, sich als kompetent zu erleben, was wiederum deren Selbstkonzept, Selbstwirksamkeitsüberzeugung und Lernmotivation fördert (Hänze et al. 2007, S. 200–205). Leistungsstärkere Schüler hingegen profitieren von der authentischen und kaum reduzierten Komplexität der Aufgabe (Kalyuga et al. 2003, S. 25). Für sie bedeuten die Hilfen ein Netz mit doppeltem Boden, denn sie können sowohl ihre Antwort ohne Hilfe der Lehrer kontrollieren als auch Gebrauch von den Hilfen machen, ohne dabei stigmatisiert zu werden. Ebenso die leistungsschwächeren Schüler. Als ergänzende Maßnahme für Leistungsstärkere und vermutlich auch schnellere Schüler bieten sich im Anschluss offenere Aufgabenstellungen an, die sich in entsprechenden Themenbereichen bewegen und Schülerinteressen bedienen (Aufschnaiter 2003, S. 22). In dieser Arbeit werden die Hilfen nach Stäudel (2008) konzipiert (siehe 9.2.8 auf Seite 248).

Zusammenfassung

Aufgaben nehmen aus didaktischer Sicht nicht nur in der Biologie eine zentrale Rolle im Unterricht ein (Seel 1981, Narciss und Proske 2001, Bruder 2003, Blömeke et al. 2006, Kuhn und Müller 2007, Thonhauser 2008). Folgende wichtige Aspekte von Aufgaben können zusammenfassend genannt werden.

Die Unterscheidung von Aufgaben kann sowohl über offensichtliche Merkmale wie die Sozialform geschehen oder über eine Typisierung von Aufgaben durch die Zusammensetzung ihrer Grundeinheiten erfolgen (Schmiemann 2013). Gerade im schulischen Kontext gilt es dabei immer zu bedenken, welche Funktion die Aufgabe im Unterricht bedient und in welcher Phase des Unterrichts sie eingesetzt werden soll (Leuders 2009). Die wichtigste Differenzierung ist die strikte Trennung zwischen Lern- und Leistungssituationen, also zwischen Lern- und Testaufgaben (Müller und Helmke 2008). Beide Formen folgen ihren eigenen Prinzipien und haben dementsprechend auch unterschiedlichen Einfluss auf das Lernen und das Unterrichtsgeschehen. Im Gegensatz zu Testaufgaben stehen Lernaufgaben am Anfang von Lernprozessen und müssen ganz anderen Qualitätskriterien entsprechen. So sollen sie relevante Inhalte transportieren und den Bedürfnissen von Schülern entsprechen, sie sollen weder unter- noch überfordern, zum Denken anregen, authentisch sein und problemorientiertes Handeln fördern, dabei am besten fachliche Kommunikation sowie soziale Interaktion begünstigen und Potenzial zur inneren Differenzierung lassen (Leisen 2010).

Innere Differenzierung kann im Unterricht nur durch Aufgaben geschehen (Schaal und Weitzel 2012). Dies kann auf folgende Arten erfolgen, die bei der Konzeption der Arbeit berücksichtigt wurden:

1. Über mehrere, unterschiedlich schwere Aufgaben kann differenziert werden. Hierbei erfolgt die Staffelung der Aufgaben entsprechend ihres Schwierigkeitsgrades (Prediger 2008, Astleitner 2008, Bohl 2012).
2. Durch entsprechende Unterstützungsmaßnahmen kann innerhalb einer einzelnen Aufgabe differenziert werden. Die Differenzierung innerhalb einer Aufgabe hat den Vorteil, dass das fachliche Niveau nicht stark gesenkt und die Kluft zwischen Schülern in Bezug auf das Wissen und Können nicht vergrößert wird. Hierfür bieten sich besonders Aufgaben mit gestuften Hilfen an (Stäudel 2009a). Aufgaben mit gestuften Hilfen besitzen zwar einen relativ hohen Schwierigkeitsgrad, sorgen durch ihre eingebauten Hilfen gleichzeitig aber dafür, dass kein Schüler auf der Strecke bleibt. Der Nachteil dieser Aufgaben ist eine sehr aufwendige Konstruktion (siehe 9.2.8 ab Seite 248).

4 ERARBEITUNG VON MATERIALIEN IN KOOPERATION MIT LEHRKRÄFTEN

4.1 Einleitung

Um empirisch validierte und praxistauglich Unterrichtsmaterialien zu erstellen, wurde eine Lehrerfortbildung durchgeführt. Eines der Ziele der Lehrerfortbildung war es, das Erfahrungswissen der Lehrkräfte zur Binnendifferenzierung zu nutzen und Hinweise zu bekommen, welche Unterstützung sich Lehrkräfte für den Schulalltag wünschen.

In den folgenden Kapiteln wird zuerst der theoretische Hintergrund der Fortbildung dargestellt und die zugrundeliegenden Theorien für das Gelingen von partizipativer Arbeit genannt. Bei Punkt 4.2 erfolgt die Beschreibung des methodischen Vorgehens bei der Erhebung der Lehrermeinungen zum Thema Binnendifferenzierung. Dieser Punkt beinhaltet sowohl die Beschreibung der teilnehmenden Personen als auch den Ablauf der Fortbildung. Ferner erfolgt eine allgemeine Beschreibung der Methode *Teilnehmende Beobachtung*. Ausgehend von der Methodenbeschreibung wird im Anschluss die konkrete Umsetzung der Methode beschrieben sowie spezifische Vorüberlegungen und Probleme genannt. Unter Punkt 4.3 werden die Ergebnisse der Lehrerfortbildung aufgeführt.

Möglichkeiten der Implementation von Maßnahmen in der Schule

In Deutschland gibt es grundsätzlich zwei Strategien, um neue Methoden in der Schule einzubinden. Zum einen gibt es die Top-Down-Strategie. Hierbei werden verbindliche Regelungen durch die Bildungsministerien vorgegeben, die dann in der Schule angewendet werden müssen. Zum anderen gibt es nach Gräsel Kooperationsverfahren, sogenannte *symbiotische Strategien der Umsetzung* (Gräsel und Parchmann 2004). Die symbiotische Umsetzungsstrategie basiert auf einer langfristigen Zusammenarbeit von Experten durch die absichtliche Bildung von *Lerngemeinschaften*, die sich aus Lehrern und Bildungsforschern zusammensetzen. In symbiotischen Settings arbeiten beide Parteien gemeinsam an einem spezifischen Problem. Der Gewinn der Zusammenarbeit entsteht durch die Einbringung der Berufserfahrung der Lehrer in Kombination mit den theoretischen Kenntnissen der Bildungsforscher (Gräsel et al. 2004). Solche Lerngemeinschaften wurden in Deutschland bereits erfolgreich durchgeführt (siehe Chemie im Kontext (Fey et al. 2004) oder Biologie im Kontext (Bayrhuber et al. 2007)). Diese Studie folgt der symbiotischen Strategie. Studienspezifische Abläufe und Interaktion zwischen den beteiligten Parteien und deren relevante Ergebnisse werden in den folgenden Abschnitten beschrieben.

Voraussetzungen für das Gelingen partizipativer Arbeit

Für einen offenen und gewinnbringenden Umgang aller Beteiligten galten sowohl bei der Planung der Fortbildung als auch bei der Durchführung Prämissen und Voraussetzungen, die den Merkmalen wirksamer Lehrerfortbildungen nach Lipowsky (siehe Tab. 15) und den Vorgaben für symbiotische Settings nach Gräsel entsprachen.

Tab. 15: Merkmale wirksamer Lehrerfortbildungen nach Lipowsky (Lipowsky und Rzejak 2012, S. 5 ff.) und deren Umsetzung in dieser Studie.

Merkmale von Wirksamkeit	Umsetzung in dieser Studie
Die Dauer der Fortbildung (keine one-shot-Veranstaltung)	10 ganz- bzw. halbtägige Treffen über 1,5 Jahre verteilt
Vertiefung von fachdidaktischem und diagnostischem Lehrerwissen mit Fokus auf Lernprozesse von Schülern	Fachvorträge mit anschließender Diskussion zu thematischen Schwerpunkten...
Verschränkung von Input-, Erprobungs- und Reflexionsphasen	...und anschließender Durchführung von praktischen Übungen in der Gruppe
Die Wirkungen eigenen Handelns erfahrbar machen	Erprobung von eigenen Materialien im Unterricht mit Weitergabe an Kollegen
Orientierung an Merkmalen lernwirksamen Unterrichts	Klare Strukturierung der Fortbildung, Mitbestimmung, lernförderliches Sozialklima ...
Feedback an die Lehrpersonen	Kollegialer Austausch innerhalb der Fortbildung zwischen allen Beteiligten

Eines der Ziele der Fortbildung war der direkte Austausch zwischen dem theoretischen Wissen der Fachdidaktiker und dem Erfahrungswissen der Lehrer. Der Dialog sollte dabei immer auf gleicher Augenhöhe stattfinden. Um diesen Austausch schon bei der Planung aktiv zu fördern, wurden nur drei Veranstaltungen im Voraus vorbereitet, die restlichen Treffen sollten explizit von den Teilnehmenden mitgestaltet und geplant werden. Für die daraus entstehenden Diskussionen und Meinungsverschiedenheiten galt es, die Konsensfindung gegenüber der Mehrheitsabstimmung zu bevorzugen. Das bedeutete, dass intensiv nach einer befriedigenden Lösung für alle Beteiligten gesucht wurde, bevor Personen von der Gruppe überstimmt wurden. Für ein solches Vorgehen war die Gruppengröße von Bedeutung, weshalb bewusst nicht mehr als 15 Personen zur Fortbildung zugelassen wurden. Ein weiterer Punkt zur Unterstützung des offenen Austausches war die Abwesenheit von offiziellen Vertretern aus Schulämtern und anderen Behörden. Um aus dem Umfeld Schule auszubrechen und Dinge mit Abstand zu betrachten, erfolgten die Treffen bewusst nicht an einer Schule, sondern immer in den Räumen der Hochschule. Darüber hinaus wurde von Anfang an auf Transparenz geachtet. Ferner wurde dargestellt, dass Forschungsinteresse sowohl an den Lehrenden als auch an deren Kooperation in der Entwicklung und späteren Erprobung der resultierenden Ansätze zur Binnendifferenzierung bestand. Für die Rekrutierung der Lehr-

kräfte galt neben der maximalen Anzahl an Teilnehmern noch ein hohes Maß an Zuverlässigkeit, Engagement und Interesse am Thema. Dies wurde dadurch gewährleistet, dass auf allen Ausschreibungen bewusst die Fortbildungsdauer von 18 Monaten ersichtlich war und sich potentielle Interessenten dessen von Anfang an im Klaren waren.

Des Weiteren ist auf die Teilnahme einer homogenen, schulartspezifischen Lehrerschaft aus folgendem Grund verzichtet worden: In Bezug auf Binnendifferenzierung und einem erweiterten Erfahrungswissen wurde bei der Anmeldung der Teilnehmenden bewusst darauf geachtet, dass sich die Gruppe aus Lehrkräften der Haupt-, Werkreal-, Realschule und Gymnasium zusammensetzte, wohlwissentlich, dass dies zu Schwierigkeiten führen könnte, die nichts mit dem Thema zu tun haben. Für die Gewährleistung der Verbindlichkeit und Nachhaltigkeit sollten zwischen den Fortbildungstagen nicht weniger als vier und nicht mehr als zehn Wochen liegen. Abb. 17 fasst das Vorhaben der Lehrerfortbildung nochmals graphisch zusammen.

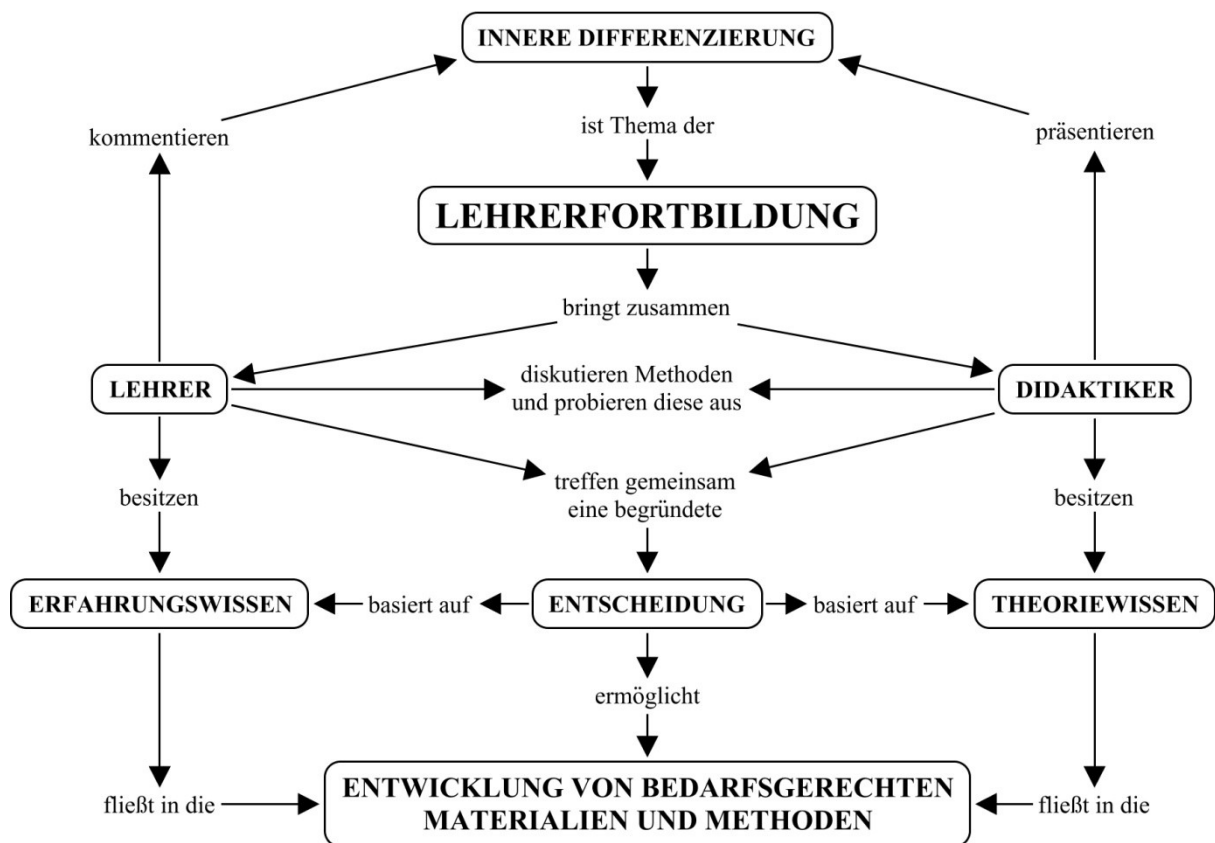


Abb. 17: Materialentwicklung zu innerer Differenzierung im Rahmen einer Lehrerfortbildung.

4.2 Methodisches Vorgehen bei der Lehrerfortbildung

Beschreibung der Teilnehmenden

Im Anschluss an die Genehmigung der Studie wurde die Fortbildung über das staatliche Schulamt Offenburg und durch das Zentrum für Weiterbildung der Pädagogischen Hochschule Freiburg beworben (Ausschreibung der Lehrerfortbildung siehe Anhang Seite 252). Nach einer mehrmonatigen Ausschreibungs- und Anmeldungsphase hatten sich insgesamt 12 Lehrkräfte angemeldet (12 ♀, 0 ♂).

Am 16. März 2010 hat das erste Treffen der Lehrerfortbildung stattgefunden. Nach der Begrüßung und Vorstellung des groben Ablaufs der Lehrerfortbildung erhielten die Teilnehmenden die Möglichkeit, sich vorzustellen. Hierbei wurde eine hohe Heterogenität innerhalb der Gruppe deutlich, was sich nicht nur der Zusammensetzung der Schularten (Lehrkräfte von der Haupt- und Werkrealschule (n=2), von der Realschule (n=6) und vom Gymnasium (n=4)) zeigte, sondern auch in den jeweiligen Biographien. Die Biographien der Teilnehmenden reichten von Junglehrern (n=2) zu Kollegen mit jahrzehntelanger Berufserfahrung (n=7), von Quereinsteigern (n=1) zu Lehrern mit 10jähriger Berufsunterbrechung (n=1) bis hin zu Schulleitern (n=1). Unter den Lehrkräften gab es Teilnehmende, die schon immer Lehrer werden wollten, bis zu Teilnehmern, die nie gedacht hätten, dass sie jemals Lehrer werden würden.

Nach den ersten Diskussionen im Plenum und Austausch von Vorerfahrungen zum Thema in Kleingruppen wurde klar, dass eine sehr große fachdidaktische und fachwissenschaftliche Heterogenität in der Gruppe bestand. Dies beeinträchtigte die konstruktive Zusammenarbeit nicht, vielmehr zeigten die Lehrkräfte großes Interesse an den Problemen, Methoden und Meinungen der Kollegen anderer Schularten. Zum Ende des ersten Treffens und zur Planung der nächsten Treffen sollten die Lehrenden ihre Erwartungen an die Fortbildung schriftlich äußern. Diese lauteten:

- Es sollte bei der Fortbildung nicht (schon wieder) um Diagnose-Vergleichs-Arbeiten (DVA) gehen (n=3).
- Die Teilnehmenden möchten ihre Ansichten in die Fortbildung einbringen und sich aktiv beteiligen (n=2).
- Die Fortbildung sollte sich auf den schulpraktischen Alltag der Lehrkräfte beziehen und auch anwendbar sein (n=5).
- Sie möchten mehr praktische Methoden zu Binnendifferenzierung lernen (n=6).
- Es sollte (ergänzende) Literatur zum späteren Nachschlagen und Weitergeben an die Kollegen geben (n=2).

Ablauf und Inhalte der Lehrerfortbildung

Unter Berücksichtigung der Erwartungen der Teilnehmenden an die Fortbildung und dem Forschungsinteresse wurde sich im Diskurs auf einen allgemeinen Ablauf geeinigt, nach dem alle Sitzungen annähernd gleich gestaltet waren. Hierbei wurde zu Beginn jeden Treffens ein theoretisch fundierter fachwissenschaftlicher oder fachdidaktischer Vortrag zu einem Thema gehalten. Der Inhalt des Vortrages wurde im Anschluss in der Gruppe auf dessen Umsetzung und Machbarkeit in der Praxis diskutiert. Die daraus resultierenden Kriterien wurden gesammelt und dienten als Grundlage für die darauffolgende Gruppenarbeitsphase. Ziel der produktorientierten Arbeitsphase war die Erstellung von Materialien, die direkt in der Praxis eingesetzt werden konnten. Diese Materialien sollten von den Lehrenden selbst bis zum nächsten Treffen in der Schule ausprobiert oder an ihre Kollegen in der Schule weitergegeben werden. Die Hausaufgabe bestand darin, eine Rückmeldung über Stärken und Schwächen des Materials in der darauffolgenden Sitzung zu geben, die dann in die Überarbeitung eingeflossen sind. Da für die Forschungsarbeit nur die Ergebnisse der Lehrerfortbildung relevant sind und nicht deren Inhalte, sind in Tab. 16 die von den Lehrenden gewünschten Themenschwerpunkte tabellarisch aufgeführt.

Tab. 16: Themenschwerpunkte der Lehrerfortbildung.

Datum/Dauer	Themenschwerpunkte
16.03.2010 halber Tag	Einführung zu Lernaufgaben, Ist-Stand und Mängelanalyse (SINUS) an bestehenden (Lern-)Aufgaben und Unterricht
13.04.2010 ganzer Tag	Differenzierungsmöglichkeiten im Unterricht, Vorstellung unterschiedlicher Aufgabentypen und Arbeit an eigenen Aufgaben zum Thema „Auge“
16.06.2010 ganzer Tag	Präsentation und Erarbeitung des Themas Auge – Vorstellung von Möglichkeiten der inneren Differenzierung und Strukturierung von Unterricht durch Aufgaben
28.09.2010 ganzer Tag	Konstruktion von differenzierenden Aufgaben zum Thema Auge. Bearbeiten und modifizieren einer ausgewählten Aufgabe
18.10.2010 halber Tag	Differenzierung in der eigenen Schule durch Schulcurriculum. Vorstellung und Einarbeitung in die Immunbiologie
10.11.2010 ganzer Tag	Einbettung von Fachinhalten in eine logische Sequenz – Planung von Unterrichtseinheiten am Beispiel der Immunbiologie (Schulcurriculum)
02.02.2011 halber Tag	Differenzierung durch den Schwierigkeitsgrad und durch die Anforderungs- und Kompetenzbereiche des Bildungsplanes mit Hilfe von Operatoren
24.05.2011 ganzer Tag	Differenzierung mit Klassenarbeiten. Differenzierung durch Noten und Erstellung eines fachspezifischen Kompetenzrasters zur individuellen Rückmeldung und Beratung
28.06.2011 halber Tag	Erstellen von Aufgaben mit gestuften Hilfen
22.11.2011 ganzer Tag	Ein- und Zuordnung der Fortbildung in das BBBB-Konzept des Kultusministeriums Baden-Württemberg, Rückblick und Verabschiedung

Allgemeine Beschreibung der teilnehmenden Beobachtung

In Anbetracht der Forschungsfrage, warum Binnendifferenzierung in der Schule bisher kaum stattfindet, fiel die Wahl der Erhebungsmethode auf die teilnehmende Beobachtung. Diese Methode dient der Erklärung von (abweichendem) Verhalten von Personen, das durch gezielte Beobachtung und Beschreibung erfolgt. Das Ziel hierbei ist die möglichst ganzheitliche Erfassung von tatsächlichem Handeln konkreter Personen in realen Situationen. Im Vergleich zu anderen Methoden liegen die Vorteile der teilnehmenden Beobachtung in der Theorie bisher Unbedachtetes aufzuspüren beziehungsweise unbekannte Facetten eines Sachverhaltes zu entdecken. Ferner gelingt es durch diese Methode, komplexe und hochgradig an einen Kontext gebundene Daten zu erheben sowie einen Einblick in die alltäglichen Routinen von Personen zu erhalten. Des Weiteren lässt die teilnehmende Beobachtung im Gegensatz zu anderen Methoden auch widersprüchliche Daten zu, aus denen sich wieder neue Forschungsfragen ableiten lassen (Klein und Pollmann 2007).

Das besondere Kennzeichen dieser Methode liegt in der persönlichen Teilnahme des Forschers an den Handlungen und sozialen Interaktionen der zu beobachtenden Personen über einen Zeitraum von möglichst einem Jahr (Goffman 1972). Für den Wissenschaftler bedeutet dies einen ständigen Wechsel zwischen Distanz und Nähe zu den teilnehmenden Personen. Ferner erfordert es sowohl ein passendes und glaubwürdiges Auftreten als auch eine gewisse Flexibilität für unvorhersehbare Situationen. Zudem muss der Forscher dezent an Gruppenprozessen teilnehmen, sich gleichzeitig aber auch ausklinken können, um den Prozess nicht zu sehr zu beeinflussen (Linden 2007).

Studienspezifische Vorüberlegungen bei der teilnehmenden Beobachtung

Wie für jede wissenschaftliche Methode müssen auch bei der teilnehmenden Beobachtung bestimmte Entscheidungen im Vorfeld getroffen werden. Hierbei unterscheiden sich nach Reißmüller Beobachtungsstudien nach dem Grad ihrer Standardisierung, also zwischen der Vorabfestlegung von konkreten Beobachtungskriterien oder dem offenen Zugang, der zwar die gleiche Forschungsfrage verfolgt, jedoch mehr Spielraum und damit mehr Interpretation und Subjektivität zulässt (Reißmüller 2009, S. 6). Des Weiteren spielt die Auswahl der Personen für die teilnehmende Beobachtung eine zentrale Rolle, da die Stichprobe unmittelbar mit der Forschungsfrage verknüpft ist (Schöne 2003, S. 10). Die Auswahlkriterien für die Testgruppe und deren Beschreibung befinden sich auf Seite 105 ff.

Nach Zimmermann (Zimmermann 2006) müssen im Vorfeld folgende Punkte geklärt werden: Was soll genau beobachtet werden? In welchem räumlichen Kontext soll die

Beobachtung stattfinden? In welchen Zeiträumen soll beobachtet werden? Wie sollen die Beobachtungsergebnisse aufgezeichnet, bzw. protokolliert werden?

Auch die Rolle des Forschers im Geschehen und im Umgang mit den Teilnehmern muss im Vorfeld definiert werden. In Abhängigkeit zur Forschungsfrage des Kontextes und der Situation sollte zwischen den folgenden Rollen unterschieden werden: Der Forscher kann zu einem vollwertigen Teilnehmer werden, sozusagen mit der Gruppe verschmelzen und alles mit machen, was die Gruppe macht. Zum anderen kann der Forscher zwar teilnehmen, sich aber ab und zu zurücknehmen und beobachten oder sich komplett aus dem Geschehen heraushalten und nur die Gruppe beobachten (Gold 1958). Ferner ist es wichtig, zu bedenken ob die Beobachtung verdeckt oder offen geschehen soll, ob die Personen also wissen, dass sie beobachtet werden oder nicht (Friedrichs und Lüdtker 1971). Abschließend muss geklärt werden, wie die Dokumentation stattfinden soll, ob sie zum Beispiel unmittelbar durch Video- oder Audio-mitschnitte aufgezeichnet wird oder der Forscher Feldnotizen, Beobachtungs- oder Gedächtnisprotokolle anfertigt (Lüders 2006).

Probleme bei der Durchführung der teilnehmenden Beobachtung

Neben der Einhaltung der Rollen, der Wahrung der Distanz und der Beachtung der strikten Neutralität können bei der teilnehmenden Beobachtung Probleme auftreten, die bei anderen Methoden eine geringere Rolle spielen. Zum einen sind das die Tagesform des Beobachters und zum anderen die steigende Müdigkeit bei zunehmender Dauer der Beobachtung. Je unkonzentrierter oder erschöpfter ein Beobachter ist, desto geringer und lückenhafter fällt die Dokumentation aus (Schöne 2003, S. 16). Das ist besonders negativ, wenn keine mediale Aufzeichnung stattgefunden hat, denn die fruchtbaren Momente in Gruppenprozessen sind nicht wiederholbar und damit schlichtweg verloren, wenn sie nicht durch den Forscher wahrgenommen wurden.

Aufgrund der langfristigen Begleitung und intensiven Auseinandersetzung mit einem überschaubaren Kreis von Personen durch die Methode der teilnehmenden Beobachtung entsteht darüber hinaus ein weiteres und nicht zu vernachlässigendes Problem. Es handelt sich hierbei um das Spannungsfeld zwischen dem zunehmenden Vertrauen der Teilnehmenden zum Forscher und dem Forschungsinteresse und der damit verbundenen Veröffentlichung der Daten. Bei der teilnehmenden Beobachtung können hoch brisante personenbezogene Daten entstehen, deren Veröffentlichung in der Beziehung zwischen Forscher und Teilnehmer einem Verrat gleich kommen. Aus diesem Grund spielt der Schutz der Identität der Teilnehmenden bei dieser Methode eine besonders große Rolle (Lüders 2006).

Konkretes Vorgehen bei der teilnehmenden Beobachtung

In Anlehnung an den explorativen Charakter wurde für diese Studie ein offener Zugang zur Beobachtung gewählt. Hierbei sollte die Auseinandersetzung der Lehrenden mit innerer Differenzierung beobachtet werden. Diese beinhaltete sowohl die theoretische Auseinandersetzung mit dem Thema beim Austausch mit den Fortbildungsleitern als auch bei Kleingruppendiskussionen untereinander. Ferner sollte auch eine praktische Auseinandersetzung der Lehrer beim Umgang und der Erstellung von eigenen Materialien beobachtet werden. Sowohl der Kontext als auch der Zeitraum der Beobachtungen bezog sich ausschließlich auf die Treffen innerhalb der Lehrerfortbildung, beinhaltete aber auch relevante informelle Unterhaltungen während der Kaffee- und Mittagspausen. Den Teilnehmenden wurde von Anfang an mitgeteilt, dass sie Teil des Forschungsinteresses sind, von daher wussten sie, dass sie auch beobachtet wurden. Die Rolle des Forschers bei dieser Fortbildung war, an den Aktivitäten der Lehrer teilzunehmen und die Lehrkräfte dabei zu beobachten. Die Dokumentation erfolgte entweder parallel oder zeitnah zu den Treffen über die Anfertigung von Feldnotizen, Beobachtungs- oder Gedächtnisprotokollen. Zur Erhöhung der Reliabilität der Daten

wurden die Beobachtungen von zwei Personen durchgeführt und anschließend miteinander verglichen (Schöne 2003).

Datenschutz (Lehrerfortbildung)

Von Beginn an wurde den teilnehmenden Lehrkräften mitgeteilt, dass Forschungsinteresse an der Zusammenarbeit besteht und daher Teile davon veröffentlicht werden. Unter der Voraussetzung der Wahrung ihrer Anonymität stimmten alle Lehrkräfte in der ersten Sitzung (16.03.2010) der Veröffentlichung einstimmig zu. Aufgrund der Wahrung der Anonymität wurden weder auf den Gedächtnis- noch auf den Sitzungsprotokollen Namen genannt oder personenbezogene Beschreibungen wie Alter, Dienstjahr oder Schulart vermerkt. Das bezieht sich auch auf die Darstellung der Ergebnisse in dieser Studie. Über die Nennung der Anzahl (n) werden keine weiteren Angaben zu den betreffenden Personen gegeben.

4.3 Ergebnisse der Lehrerfortbildung

In den folgenden Abschnitten werden die Resultate der Lehrerfortbildung aufgeführt. Nach der Beschreibung der primären Datenerhebungsmethode wird zuerst die Rückmeldung der Lehrer aufgeführt, die die Qualität der gesamten Fortbildung aus Teilnehmersicht beschreibt. Entsprechend der Forschungsfragen wird im darauffolgenden Abschnitt aus Lehrersicht begründet, warum Binnendifferenzierung in der Schule bisher kaum stattfindet. Basierend auf diesen Begründungen und den erfahrungsbasierten Ansprüchen der Lehrer an eine Differenzierungsmaßnahme in der Schulpraxis werden im letzten Abschnitt die Wünsche der Lehrkräfte aufgeführt, die schließlich die Basis für die kooperative Entwicklung konkreter Unterrichtsmaterialien bilden.

4.3.1 Warum findet Binnendifferenzierung bisher kaum statt?

Auf die Frage, warum Binnendifferenzierung bisher kaum im alltäglichen Unterricht stattfindet, führten die Lehrenden über die Treffen und Themen hinweg mehrere Gründe auf. Der häufigste Grund hierbei war Ratlosigkeit. Dies wurde darin deutlich, dass Lehrende zwar den Sinn und Nutzen der Anpassung von Unterricht an die Bedürfnisse der Lernenden kennen, allerdings nicht wissen, wie das funktionieren soll. Folgende Frage bringt das Problem auf den Punkt: *Wenn ich zum ersten Mal differenziere, wie mache ich das?* Dieses grundsätzliche Problem zeigte sich auch darin, dass es nach Lehrermeinung zu wenig konkrete Startvorgaben und Beispiele für die Unterrichtspraxis gibt, ganz zu schweigen von gutem Unterrichtsmaterial. Auch beginnt und endet die Vorbereitung von Unterricht oft mit den Schulbüchern, jedoch zeigen diese meist keine Ansätze für innere Differenzierung. Des Weiteren stehen Lehrende der Masse und Anzahl an Inhalten ratlos gegenüber, die am besten alle und noch besser auf

verschiedenen Wegen vermittelt werden sollen. Als weiteres Problem bei den Inhalten sehen Lehrer die eigene Fachkompetenz als Problem. So haben sie nach eigener Aussage nicht bei allen Themen den ausreichenden Überblick oder die Tiefe an Fachwissen, um eigenständig Inhalte zu streichen oder im Sinn der Unterrichts Anpassung entsprechend zu organisieren, ganz zu schweigen von den Fächern, die sie fachfremd unterrichten müssen.

Ein großes Problem sehen Lehrer auch in den Rahmenbedingungen von Unterricht. Wie sorgt man bei der Einführung von Differenzierung und den damit verbundenen Konsequenzen für Akzeptanz bei den Schülern, Eltern und Kollegen? So äußerte eine Person, dass manche Schüler Differenzierung als Strafe empfunden hätten und manche Eltern es nicht fair empfanden, wenn Schüler nicht das Gleiche an Aufgaben bekommen haben. Auch wurde gesagt, dass im Lehrerkollegium Zweifel an der Machbarkeit von innerer Differenzierung bestehen, denn das Aufbrechen von bewährten Unterrichtsabläufen und Strukturen kann leicht zu Problemen bei der Organisation von Unterricht und damit bei der Disziplin führen. Die Vorgabe von mindestens vier schriftlichen Klassenarbeiten und Vergleichsarbeiten in den Kernfächern in Baden-Württemberg mache innere Differenzierung schwierig. Es sei paradox, denn wie soll man Schüler unterschiedlich behandeln, wenn man diese am Ende wieder auf gleiche Art und Weise prüft? So wird bei der Forderung der Wissenschaftler nach individualisiertem Unterricht die Notengebung nicht berücksichtigt. Ferner fehlen den Lehrern in der Lehreraus- und Weiterbildung die Möglichkeiten, notwendige diagnostische Fähigkeiten zu erlernen und zu verbessern.

Genauso problematisch wie die schulischen Rahmenbedingungen sehen Lehrende den potentiellen Mehraufwand von Binnendifferenzierung. Dieser Mehraufwand bezieht sich pro Klasse sowohl auf die Vor- und Nachbereitung von Unterricht als auch auf dessen Durchführung. Hierbei gilt zu bedenken, dass die arbeitserleichternden Parallelklassen zwar thematisch weiterhin annähernd gleich laufen, jedoch durch innere Differenzierung pro Klasse höchst wahrscheinlich anders aufbereitet werden müssen. Als Begründung für den relativ kleinen Anteil an differenzierenden Ansätzen im Unterricht bemängeln Lehrkräfte die bereits bestehende zeitliche Auslastung durch das Stundendeputat und den darüber hinausgehenden Verpflichtungen von Lehrern an den Schulen.

Bei den Diskussionen wurde ein weiterer Aspekt von Differenzierung deutlich, der in der Theorie bisher wenig zu finden ist. Nach Lehrermeinung besteht nämlich keinerlei Austausch zwischen den Schularten. Das ist nicht nur problematisch für Schüler, die die Schulart wechseln, sondern auch für Lehrende, die diese Schüler neu in die Klasse bekommen. Beispielsweise müssen sich Realschullehrkräfte mit ganz anderen Voraus-

setzungen durch neue Schüler von Gymnasien und mit Schülern von Werkrealschulen auseinandersetzen, was das Leistungsspektrum ihrer neu zusammengesetzten Klassen stark erweitert. Eine Vorbereitung oder Pufferung dieses Aufeinandertreffens ist meistens nicht möglich, da keine Informationen über bereits vermittelte Methoden und Inhalte aus der vorherigen Schule bekannt sind. Dies gilt aber nicht nur für den Wechsel der Schulart, sondern auch für den Wechsel von einem Gymnasium zu einem anderen in der gleichen (!) Stadt.

Im Kontrast zu den gerade beschriebenen Problemen von Binnendifferenzierung in der Schulpraxis, besteht bei Lehrenden ein hohes Interesse an innerer Differenzierung. Dies besteht vor allem dann, wenn an konkreten Beispielen gearbeitet wird und Lehrer nicht alleine damit dastehen, sondern sich gegenseitig Anregungen geben können. Lehrende finden nämlich, dass es trotz der Aufteilung der Schüler auf die drei Schularten teilweise enorme Leistungsunterschiede innerhalb einer Klasse gibt. Ferner geht es bei innerer Differenzierung nicht nur darum, die Schwächeren zu unterstützen, sondern auch die Stärkeren zu fördern.

Auf die Frage, was Lehrer konkret zur Verbesserung der Lage tun würden, äußerte der Großteil, dass sie die Themenvorgaben radikal entrümpeln würden. Das begründeten sie damit, dass durch die Entspannung der bisherigen Informationsdichte, auch mehr Zeit und Raum für Pädagogik und Didaktik geschaffen würde. Des Weiteren würden sie den Klassenteiler reduzieren, was automatisch zu einer besseren Betreuung von Schülern führen würde. Außerdem würden sie Sozialpädagogen und Psychologen an den Schulen einstellen, die sich dann um die Fälle kümmern könnten, die über die Möglichkeiten einer fachlichen Differenzierung hinausgehen.

4.3.2 Welche Differenzierungsmaßnahmen bevorzugen Lehrkräfte?

Nach vielen Diskussionen über Vor- und Nachteile und dem Ausprobieren verschiedener Differenzierungsansätze (siehe Abb. 6 auf Seite 22) einigten sich die Lehrer unserer Fortbildung für das Gelingen einer Differenzierungsmaßnahme auf folgende Punkte:

(1) Eine Differenzierungsmethode sollte von Lehrern geplant werden und ausdrücklich leistungsschwächere Schüler unterstützen. (2) Sie sollte die Wissenslücke zwischen leistungstärkeren und leistungsschwächeren Schülern nicht erweitern, sondern alle Schüler auf den gleichen Wissensstand bringen, damit sie am Ende der Unterrichtseinheit auch die gleiche Klassenarbeit schreiben können. (3) Entsprechende Aufgaben oder Anleitungen sowie Lösung sollten dabei klar und einfach formuliert sein. (4) Das Antwortformat sollte eng und gut definiert sein und nicht viel Zeit- und Materialaufwand von Lehrer erfordern. (5) Darüber hinaus sollten die Materialien so wenig Text

wie nötig verwenden. Den letzten Punkt unterstützten die Lehrer mit dem Argument, dass Leistungsschwäche in Biologie nicht nur auf unzureichenden Fachkenntnissen der Schüler beruht, sondern oft mit Schwächen im Verständnis und der Produktion von Texten einhergeht.

Für die praktische Umsetzung in der Schule wurde von den Lehrern darüber hinaus angemerkt, dass sich eine erfolgreiche Differenzierungsmaßnahme der Dinge bedient, die in einer normalen Schule vorhanden sind. Nach Meinung der Lehrer schließe das den Einsatz von Computern und Programmen aus (6). Denn falls PCs an der Schule überhaupt vorhanden sind, dann sind diese meist veraltet oder nicht professionell gewartet. Ferner stehen meistens zu wenige PCs und damit Arbeitsplätze zur Verfügung. Oft sind die Schüler im Umgang mit Computern geübter als ihre Lehrer und machen am PC häufig nicht das, was sie tun sollten. Aus diesem Grund scheuen sich viele Kollegen, die „neuen“ Medien einzusetzen, da es dazu führen würde, dass der Differenzierungsansatz nicht in der Schule ankommen würde. Als Vorschlag nannten die Lehrer die klassische Kopiervorlage als Material. Diese würde in die bereits bestehenden Materialordner der Lehrer integriert werden und könnte bei Bedarf einfach hervorgeholt und kopiert werden.

Basierend auf den Anforderungen der Lehrer wurde ein Differenzierungsansatz innerhalb eines partizipativen Prozesses entwickelt. Die Wahl der Gruppe fiel auf innere Differenzierung durch Lehrer zur Förderung leistungsschwächerer Schüler mit Hilfe von Experten-Concept-Maps. Der Prozess der symbiotischen Entscheidungsfindung wird in Abb. 18 dargestellt.

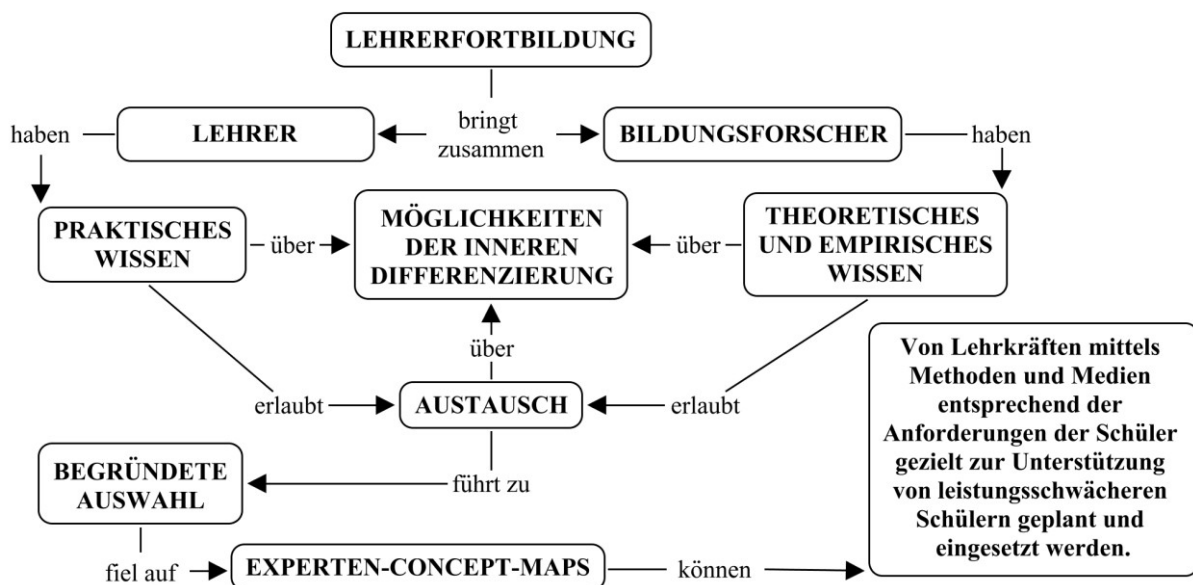


Abb. 18: Darstellung des symbiotischen Prozesses bei der Entwicklung der Differenzierungsmethode.

4.3.3 Rückmeldung der Lehrenden auf die Fortbildung

Über das Thema Differenzierung hinaus gaben die Teilnehmenden folgende wichtige Rückmeldungen zu der Atmosphäre und allgemeinen Aspekten der Durchführung (konstruktive Arbeitsphasen, Zeit zum Nachdenken):

Methodisches Vorgehen: Die Rückmeldungen wurden nach jedem Treffen auf einem standardisierten Rückmeldebogen schriftlich erhoben. Dieser beinhaltete die geschlossenen Frage, welche Note die Lehrer dieser Sitzung geben würden (von 1 bis 6) und die offene Frage, was an dieser Sitzung gut oder weniger gut war und was den Teilnehmenden gefehlt hat.

Ergebnis: Die Noten der Teilnehmenden wurden am Ende der Fortbildung zusammengezählt und der Durchschnitt ermittelt. Die Gesamtnote lautete *sehr gut* bis *gut* (Note 1,5). Die Verbalbeurteilungen wurden gesichtet, Kategorien zugeordnet und anschließend ausgezählt. Die Äußerungen der Lehrer sind im Folgenden aufgeführt.

Der stärkste Punkt für eine erfolgreiche Lehrerfortbildung war die Atmosphäre (n=7). Besonders wurde der offene Umgang der Teilnehmer untereinander als auch mit den Fortbildungsleitern betont. Hierbei ging es vor allem um den respektvollen, fachlichen Austausch, bei dem auch persönliche Ängste und Zweifel geäußert und ernst genommen wurden: *Wie findest du das? Ist das gut so wie ich das mache? Wie könnte ich es anders machen? Wie machst du das? Welche Probleme hast du?*

Folgende Punkte waren nach Lehrermeinung der Förderung dieser Offenheit besonders dienlich:

- Die simple örtliche Trennung der Treffen außerhalb einer Schule
- Der Ausbruch aus dem Mikrokosmos des eigenen Lehrerkollegiums und besonders der Ausbruch aus den eigenen Routinen bei der Unterrichtsplanung
- Durch die schulartübergreifende Zusammensetzung bei der Arbeit in kleinen Gruppen kam es durch die unterschiedlichen Perspektiven zu anregenden Diskussionen und zu einer Erweiterung der eigenen Sichtweise, die mit Kollegen der eigenen Schulart nicht möglich waren. Es wurde auch gesagt, dass gerade die unterschiedlichen Ansichten innerhalb der Gruppe zu guten Lösungen geführt haben.
- Die konstruktive Arbeit (n=6). Hierbei war ihnen besonders wichtig, dass klare Ziele gesetzt und auch erreicht wurden. Es wurde als positiv empfunden, dass die Gestaltung der Fortbildung offen war und man seine eigenen Interessen einbringen und den Verlauf mitbestimmen konnte. Dies bezog sich aber nicht nur auf die Ziele, sondern auch auf die Inhalte. Hierbei wurde zwar der stete Bezug zur Schule

gelobt, aber auch die Möglichkeit, sich mal wieder länger und ausführlicher mit wissenschaftlichen Inhalten auseinander setzen zu können, da hierfür im Alltag meist keine Zeit bliebe.

- Der Faktor Zeit (n=3) wurde auch im Zusammenhängen damit genannt, dass es als lehrreich empfunden wurde, Zeit zum Nachdenken zu haben („Das arbeitet in mir“), einen Fachinhalt mal wieder genau zu analysieren, sich Zeit für die Entwicklung einer eigenen Aufgabe zu nehmen oder einen Arbeitsauftrag noch verständlicher zu formulieren.
- In Verbindung mit der Zeit haben die Lehrenden aber auch den Ablauf der Veranstaltung gelobt. Die veranschlagten 18 Monate waren nach Lehrermeinung zwar lang, wurden aber als positiv wahrgenommen. Hierbei wurde gesagt, dass es sich durch die Dauer nicht um eine Blitzlicht-Veranstaltung handelte, sondern dass die Zeit nachhaltig war und Langzeitwirkung hatte. Dass Probleme nicht aus Zeitgründen nur angekratzt und oberflächlich behandelt wurden, sondern diesen auf den Grund gegangen und eine Lösung entwickelt wurde.
- In Bezug auf den Ablauf der Fortbildung (n=5) empfanden die Lehrer positiv, dass es durch die Häufigkeit der Termine zu einer Regelmäßigkeit und Routine kam, wobei die frühzeitige Bekanntgabe der Termine den Lehrern die Zeitplanung erheblich erleichterte.
- Als abschließenden und nicht zu unterschätzenden Punkt wurde erwähnt, dass nicht nur für Verpflegung der Teilnehmer gesorgt wurde, sondern auch dass Pausen eingeplant und vor allem auch eingehalten wurden.

4.4 Zusammenfassung

Ausgehend von dem Problem, dass zwar theoretisch begründete Ansätze für Binnendifferenzierung existieren, diese jedoch kaum in der Schulpraxis umgesetzt werden, galt es, zwei Fragen zu beantworten: Warum ist das so und was kann dagegen unternommen werden? Zur Beantwortung wurden Lehrer aller Schularten zu einer 18-monatigen Fortbildung eingeladen. Die Fortbildung entsprach einem symbiotischen Setting (Gräsel und Parchmann 2004) und den Merkmalen wirksamer Lehrerfortbildungen (Lipowsky und Rzejak 2012). Die Gespräche wurden mittels teilnehmender Beobachtung sowie Feldnotizen und Gedächtnisprotokollen dokumentiert und qualitativ ausgewertet (Schöne 2003).

Die Auswertung hat gezeigt, dass Lehrer der Thematik der inneren Differenzierung oft ratlos gegenüberstehen. Das liegt daran, dass es nicht nur an konkreten Vorgaben mangelt, sondern auch an persönlichem Fachwissen und diagnostischer Erfahrung. Neben den eigenen Hemmungen sind aber auch schulische Rahmenbedingungen bei der Durchführung hinderlich. Ferner besteht bei innerer Differenzierung häufig eine mangelnde Akzeptanz bei Eltern, Schülern und Kollegen. Darüber hinaus besteht die Angst vor unverhältnismäßigem Mehraufwand und daraus resultierender Überfrachtung für Lehrer im Schulalltag.

Zur Verbesserung des Gelingens von Binnendifferenzierung in der Schule schlugen die Lehrer vor, die Themenvorgaben zu entrümpeln und damit die Informationsdichte zu verkleinern wodurch Zeit und Raum für Kompetenzentwicklung geschaffen wäre. Für eine bessere Betreuung der Schüler sollte der Klassenteiler reduziert werden sowie Sozialpädagogen und Psychologen als Unterstützung an die Schulen kommen. Diese sollten sich um jene Probleme bei Schüler kümmern, die über die Kompetenz eines Fach- oder Klassenlehrers hinausgehen.

Auf die Frage, was eine Differenzierungsmaßnahme für den Einsatz in der Praxis ausmache, äußerten die Lehrer, dass sie leicht in der Anwendung sein muss und keinen großen Mehraufwand bei der Vor- und Nachbereitung von Unterricht erfordern darf. Ferner soll sie gezielt leistungsschwächere Schüler unterstützen. Darüber hinaus sollte sie von Lehrern geplant und zugeteilt werden, da die leistungsbezogene Selbsteinschätzung gerade junge und ungeübte Schüler zu Beginn überfordert und geübt werden muss. Des Weiteren wurde explizit gefordert, dass die Methode das fachliche Niveau nicht weiter senken soll und die Kluft zwischen den Schülern nicht erweitern darf, damit alle Schüler am Ende einer Unterrichteinheit auch die gleiche Klassenarbeit schreiben können.

5 ALLGEMEINE METHODISCHE VORGEHENSWEISEN UND EINGESETZTE MATERIALIEN

5.1 Einleitung

Ausgehend von dem Mangel an innerer Differenzierung im Biologieunterricht (siehe Seite 1) und entsprechend der Ergebnisse der Forschungsfragen (a) aus der Lehrerfortbildung (siehe Seite 97 ff.) werden in den Abschnitten dieses Kapitels die zentralen Materialien und Methoden zu den folgenden Forschungsfragen vorgestellt:

- a) Welche Ansätze innerer Differenzierung eignen sich aus fachdidaktischer Sicht für die Vermittlung biologischer Inhalte? Welche dieser Ansätze nehmen Lehrkräfte aufgrund ihres Erfahrungswissens an?
- b) Hat eine differenzierende Darstellungsform annähernd identischer Lehrinhalte über Experten-Concept-Maps oder Text Auswirkungen auf das Wissen und Können bei Lernenden mit unterschiedlichen Leistungsniveaus?
- c) Treten bei den eingesetzten Differenzierungsmaßnahmen geschlechtsspezifische Unterschiede auf?
- d) Welchen Einfluss hat die Lernzeit auf das Ergebnis der Schüler?
- e) Wie gestaltet sich der Lernprozess der Schüler mit den differenzierten Darstellungsformen und der Lernaufgaben und wo treten dabei Schwierigkeiten auf?

Dies Kapitel beinhaltet die Rekrutierung und Beschreibung der teilnehmenden Schüler und die Einteilung der Schülergruppen. Ferner wird die formative Evaluation, sowie die daraus resultierten Materialien und der Ablauf der Interviews vorgestellt. Zum Ende des Kapitels werden die eingesetzten Methoden der summativen Evaluation aufgeführt, die die Grundlage des darauffolgenden Kapitels bilden (Empirische Untersuchung siehe Seite 125 ff.).

5.2 Rekrutierung und Beschreibung der Testpersonen

Bei dieser Forschungsarbeit handelte es sich um eine quasi-experimentelle Studie im 2x2-Design. Quasiexperimentell bedeutet hierbei, dass die Versuchspersonen aufgrund der Forschungsfragen nicht randomisiert in Versuchsgruppen eingeteilt werden konnten, sondern bestimmte Kriterien für die Zuteilung erfüllen mussten (siehe unten).

In Abhängigkeit zum Bildungsplan von Baden-Württemberg (2004) und den Interessen der Lehrer, wurde das menschliche Auge zum thematischen Inhalt der inneren Differenzierung in dieser Forschungsarbeit. Die Festlegung hatte zur Folge, dass sich auf jene Klassenstufen beschränkt werden musste, die das Thema betrifft (Realschule Klasse 6 und 7). Wichtig war dabei auch, dass die betreffenden Schüler das Thema „Auge“ bisher nicht im Unterricht hatten.

Nach der Genehmigung der Forschungsarbeit durch das Kultusministerium von Baden-Württemberg (Aktenzeichen 31-zu 6499.20/642 am 14.07.2010) wurden über eine landesweite Ausschreibung (siehe 9.4 auf Seite 252) acht Lehrkräfte mit insgesamt 11 Klassen für die Teilnahme gewonnen. Im darauffolgenden Schritt wurden die Erziehungsberechtigten von 291 Schülern durch einen Brief über das Vorhaben informiert. Von diesen 291 Schülern wurde nur 148 Schülern das Einverständnis an der Teilnahme durch ihre Erziehungsberechtigten erteilt und schriftlich bestätigt. Hierbei handelte es sich ausschließlich um Realschüler aus Schulen mit eher ländlichen Einzugsgebieten aus den Landkreisen Tuttlingen (Schwarzwald) und Ortenau im Kinzigtal. Die beteiligten 148 Schüler haben im nächsten Schritt einen schriftlichen Fragebogen ausgefüllt (Details zum Fragebogen siehe Seite 219). Dieser Fragebogen verfolgte zwei Ziele: Zum einen, welche Schüler waren überhaupt für die Teilnahme an den Interviews geeignet (Ein- und Ausschlusskriterien)? Zum anderen, welche der geeigneten Schüler sich entweder einer leistungsstärkeren oder leistungsschwächeren Gruppe zuordnen ließen? Im Folgenden werden die Kriterien zur Teilnahme an der Studie beschrieben:

- Die Schüler benötigen ein ausreichendes Text- und Sprachverständnis, um die Aufgaben und den Input bearbeiten zu können, was über ausgewählte Auszüge aus dem KFT 4-12+R (Heller und Perleth 2000) gewährleistet ist. Der KFT wird auf ausgewählte Items aus dem ersten Teil des kompletten Tests beschränkt, sodass die Durchführungsdauer für den Schulalltag reduziert und somit im laufenden Schulbetrieb verwendet und von den Lehrern eingesetzt werden kann. Eigentlich war hier der Einsatz des CFT 20-R nach Weiß (2006) geplant, allerdings wurde der Test nicht zugelassen. In Baden-Württemberg gilt die Einschränkung, dass ausschließlich Psychologen Intelligenztest einsetzen und auswerten dürfen.
- Die Schüler sollten kein bis kaum Vorwissen zum Thema haben. Der Grund hierfür liegt darin, dass Schüler mit thematischem Vorwissen die Aufgaben im Interview befriedigend beantworten können. Durch das Vorwissen wären Schülerantworten unabhängig von dem eingesetzten Input richtig, was das Ergebnis verfälschen würde. Aus diesem Grund wurde das Vorwissen der Schüler zum Thema Auge über geschlossene und offene Fragen abgeprüft.

- Die Schüler müssen über ein ausreichendes Verständnis für Concept-Maps verfügen. Ein unzureichendes Verständnis für Concept-Maps würde die Testung dieses Inputs irrelevant machen, da aus einem Input, der nicht verstanden wird, auch nichts gelernt werden kann.

Die Ergebnisse aus dem Fragebogen wurden mit den betreffenden Lehrenden diskutiert und durch die Zeugnisnoten der Schüler in den Fächern Biologie, Mathematik und Deutsch ergänzt. Da es hier um die Vermittlung biologischer Inhalte geht, wurde die Biologienote als ausschlaggebender Faktor für die Zuteilung von Schülern verwendet. Die Zuteilung in die leistungsstärkere oder leistungsschwächere Gruppe erfolgte demnach ausschließlich über die Schulnote in Biologie und nicht über den Intelligenzquotienten oder Ähnliches. Um für die Berechnungen annähernd gleich große Gruppen zu erhalten, wurde die Gesamtheit aller Schüler über den Median-Split der Biologienote geteilt. Nach der Berücksichtigung der Einverständniserklärung der Erziehungsberechtigten, aller genannten Ein-, Aus- und Zuteilungskriterien und dem Überwinden von schulischen und organisatorischen Hürden¹⁴ blieben am Ende der Datenerhebung von den ursprünglich 291 Schülern eine Anzahl von insgesamt 112 Testpersonen übrig. Die Stichprobe für diese Studie setzt sich aus insgesamt 62 Mädchen und 50 Jungen zu gleichen Teilen aus den Klassen 6 und 7 zusammen. Aufgrund des ländlichen Einzugsgebietes befinden sich sehr wenige Schüler mit Migrationshintergrund in der Stichprobe.

5.3 Datenschutz (Fragebögen und Schülerinterviews)

Bei der Verwendung persönlicher Daten in Kombination mit Schulnoten ist nicht nur aus schulrechtlicher Sicht ein verstärkter Datenschutz erforderlich. Aus diesem Grund wurde das folgende Verfahren angewendet: Alle Fragebögen wurden mit einer eindeutigen ID versehen. Die teilnehmenden Lehrer fertigten eine Liste der teilnehmenden Schüler an. Auf dieser Liste ordneten sie, nach der Abgabe des Fragebogens, die jeweilige ID des Fragebogens dem entsprechenden Schüler zu. Diese Liste verblieb beim jeweils teilnehmenden Lehrer. Lediglich die Fragebögen, versehen mit der ID, wurden zur Auswertung an die Forschenden gesendet. Diese fertigten wiederum eine Liste an, die den IDs die Testdaten zuwies. Somit war den teilnehmenden Lehrern lediglich bekannt, welcher Schüler an der Studie teilgenommen hat und welcher nicht. Sie waren jedoch zu keiner Zeit in der Lage, die erbrachten Leistungen (Testdaten) mit einem Schüler in Verbindung zu bringen. Im Umkehrschluss war es den Forschenden

¹⁴ Stundenplanänderungen, Schulabgänger, Wandertage, Bundesjungendspiele, Schüleraustausch, betreffende Lehrkräfte oder Schulleitung nicht informiert, Krankheit, Gebäude im Umbau, Raum- und Zeitmangel, (un)angekündigte Tests und Klassenarbeiten, pädagogische Tage, spontane Schüchternheit von Schülern direkt vor dem Interview, Europa Park, Evaluation der Schule, Berufsorientierung an der Realschule (BORS), Gesundheitstage, Praktika, Chor- und Theaterproben, Schulfeste etc..

nicht möglich, die Testdaten und die Ergebnisse einem Schüler namentlich zuzuweisen. Diese Aufteilung der Informationen gewährleistete den Schutz der Daten aller Beteiligten.

5.4 Einteilung der Gruppen

In Kapitel 6 (siehe Seite 125 ff.) werden Schülergruppen miteinander verglichen. Alle Gruppen wurden anhand der gleichen Kriterien gebildet (siehe nachfolgende Liste). Die Berechnungen der Vergleiche fanden immer mit den gleichen Gruppen und in der gleichen Reihenfolge statt:

- 1) **Alle Schüler:** Bevor jede Gruppen miteinander verglichen wird, werden immer zuerst die Daten aller Schüler dargestellt. Die Anzahl der Schüler kann je nach Datenmaterial, Berechnungsmethode oder fehlenden Werten variieren. Sofern dies der Fall ist, wird es bei den Ergebnissen erwähnt.
- 2) **Leistungsstärkere und leistungsschwächere Schüler:** Die Gesamtheit aller 112 Schüler wurde durch den Median der Biologienote geteilt. Der Median lag bei 2,50 und teilte die Gruppe in 61 leistungsstärkere und 51 leistungsschwächere Schüler.
- 3) **Input:** 57 (n) Schüler erhielten den Fachtext und 55 (n) Schüler die Experten-Concept-Map als Input.
- 4) **Geschlecht:** In Bezug auf das Geschlecht setzt sich die Schülergruppe aus 50 Jungen und 62 Mädchen zusammen.
- 5) **MC-Ergebnis:** Nach der Auswertung des Multiple-Choice-Tests wurden durch den Median (23,50) der Ergebnisse aller Schüler zwei weitere Gruppen gebildet. Sowohl die Gruppe der Schüler mit einem guten als auch die Gruppe der Schüler mit einem schlechten Ergebnis im Multiple-Choice-Test bestehen jeweils aus 56 (n) Schülern.

Die Zahlen beziehen sich auf die Rohdaten und beinhalten sowohl die Ausreißer (Outliers) als auch die Schüler mit fehlenden Daten. Je nach Forschungsfrage und Art der Berechnung kann die Schüleranzahl im Folgenden leicht variieren.

5.5 Formative Evaluation des Input und der Aufgaben

Bevor die Wirkung der Interventionsmaßnahme zur inneren Differenzierung in der Praxis summativ überprüft werden konnte, wurden alle Materialien formativ evaluiert. Hierbei wurde besonders Wert auf die Vergleichbarkeit und Verständlichkeit der beiden Repräsentationsformen Text und Experten-Concept-Map gelegt.

Dies bedeutete, dass über die beiden Formen des Inputs hinaus auch bei den Abbildungen und den Hilfen geklärt werden musste, welche Ansprüche durch deren Einsatz in Bezug auf das Lösen der vier Aufgaben entstehen. Die Ansprüche mussten sich in allen Materialien widerspiegeln und auf alle Materialien passen. Die stetige Kontrolle aller Materialien, deren Überarbeitung und erneuter Kontrolle war ein zyklischer Prozess mit mehreren Durchgängen, der in Abb. 19 dargestellt ist und im anschließenden genauer beschrieben wird.

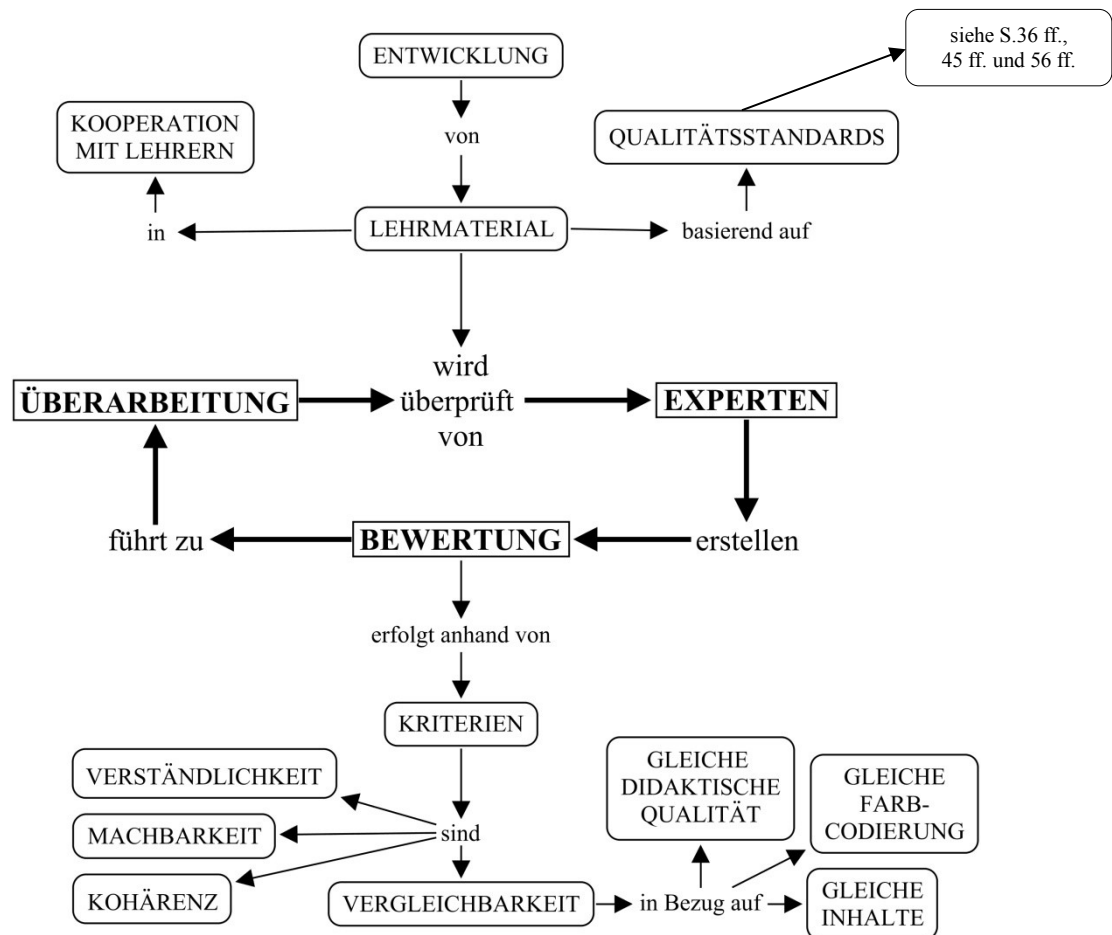


Abb. 19: Prozess der zyklischen formativen Evaluation.

Zu Beginn erfolgte die Entwicklung des Fachtextes mit passenden Inhalten für die Schulart und Klassenstufe. Der Fachtext musste in das Format der Experten-Concept-Map überführt werden und vergleichbar sein. Zeitgleich wurden die Abbildungen erstellt und an den Input angepasst. Die Entwicklung und die nach Kriterien geleitete Überprüfung erfolgten durch die Lehrer der Fortbildung (erster Durchgang). Daraufhin wurden alle Materialien von externen Personen auf Machbarkeit, Vergleichbarkeit und Verständlichkeit überprüft (zweiter Durchgang). Die zweite Überprüfung geschah über eine externe Gymnasiallehrerin, zwei Diplom-Psychologinnen und eine Sozialpädagogin. Abschließend wurden alle Materialien nochmals von 27 Lehramtsstudierenden überprüft (dritter Durchgang). Die Überprüfung durch die Lehramtsstudierenden be-

inhalten nicht nur die Begutachtung des Materials, sondern auch drei Probeinterviews. Hierbei wurden nicht nur der Ablauf und die Stimmigkeit der Materialien und Anweisungen überprüft, sondern auch verschiedene Kamerasysteme ausprobiert. Im Anschluss daran wurden die Materialien im Herbst 2011 zwei Mal mit jeweils vier Schülern erprobt und entsprechend modifiziert (vierter und fünfter Durchgang). Die im Folgenden beschriebenen Materialien beziehen sich ausschließlich auf jene, die in der Studie verwendet wurden.

5.6 Vorstellung des differenzierten Input

Im Folgenden werden der Fachtext und die Experten-Concept-Map dargestellt.

Tab. 17: Der Text als fachlicher Input (ohne Farbleitcode).

Im hinteren Teil des Auges befindet sich die Netzhaut. In der Netzhaut des Menschen befinden sich rund 130 Millionen lichtempfindliche Sehsinneszellen. Die Sehsinneszellen sind die einzigen Zellen im Körper, die Lichtreize wahrnehmen können. Sie sind über andere Nervenzellen mit Schaltzellen verbunden. Die Fortsätze der Schaltzellen ziehen in den Sehnerv. Der Sehnerv führt zum Gehirn. An der Stelle, wo der Sehnerv den Augapfel verlässt, liegt der blinde Fleck. Im Bereich des blinden Flecks sind keine Sehsinneszellen vorhanden.

Es gibt zwei Typen von Sehsinneszellen, die Stäbchen und die Zapfen. Die Stäbchen sind sehr lichtempfindlich, mit ihnen können wir in der Dämmerung und nachts sehen. Allerdings können sie nur zwischen Hell und Dunkel unterscheiden. Die äußere Form der Stäbchen ist länglich und schmal. Die Zapfen sind wenig lichtempfindlich. Im Vergleich zu den Stäbchen sind die Zapfen in ihrer Form eher kurz und dick. Mit den Zapfen können wir allerdings Farben erkennen und unterscheiden.

Die Netzhaut ist nicht an allen Stellen gleich aufgebaut. Im Randbereich der Netzhaut gibt es nur Stäbchen. Im zentraleren Bereich befinden sich Stäbchen und Zapfen. An einer Stelle im Zentrum befinden sich nur Zapfen in der Netzhaut. Hier ist jeder Zapfen über andere Nervenzellen mit einer Schaltzelle verbunden. Daher liegt dort die Stelle des schärfsten Sehens. Diese Stelle wird der gelbe Fleck genannt. Beim Fixieren (konzentriert anschauen) eines Gegenstandes wird das Auge immer so ausgerichtet, dass das Bild direkt auf den gelben Fleck in der Netzhaut trifft.

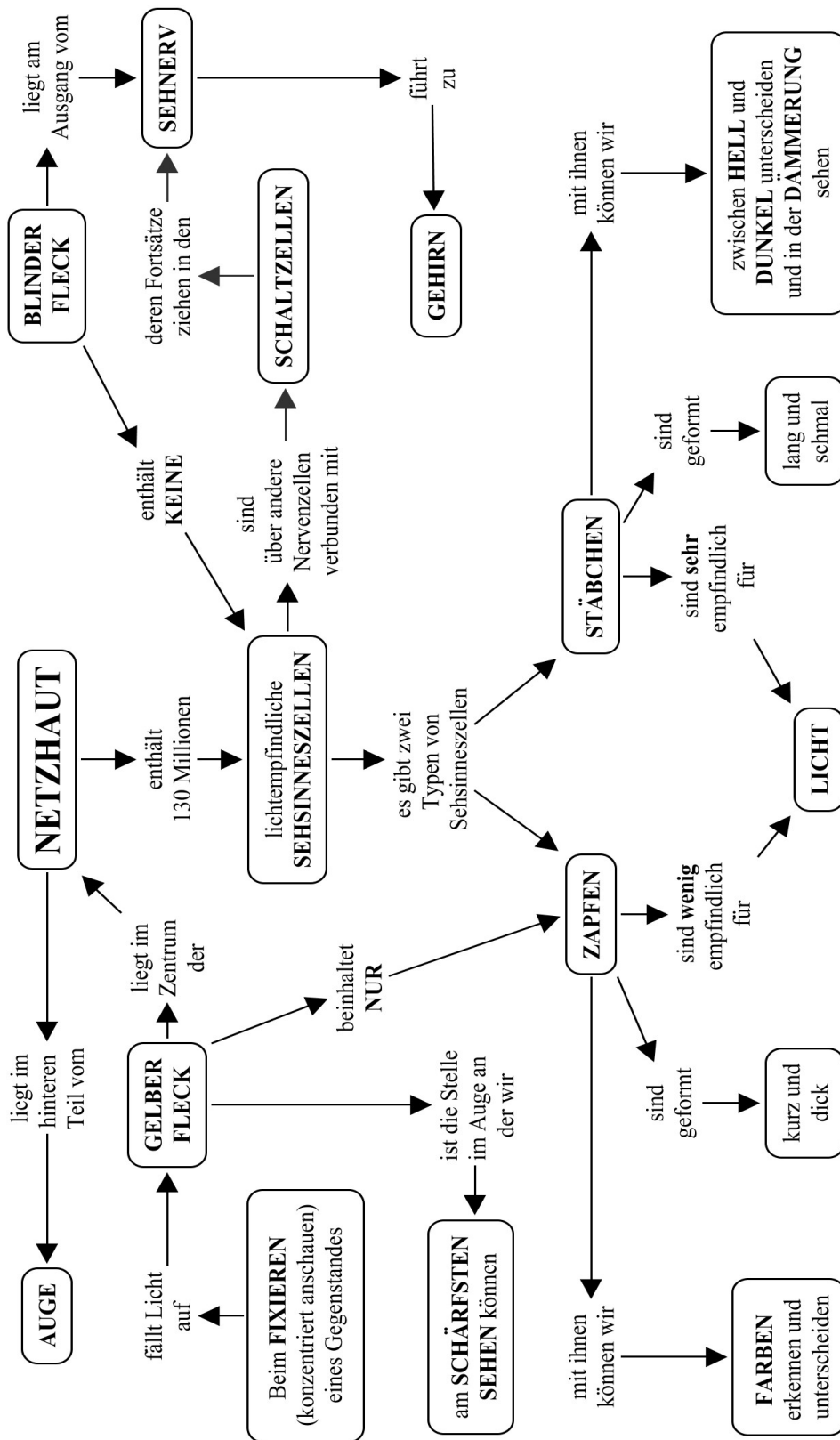


Abb. 20: Die Experten-Concept-Map als fachlicher Input (ohne Farbleitcode).

5.7 Vorstellung der Lernaufgaben 1 bis 4

Für die Formulierung der Arbeitsaufträge war es wichtig, dass die Lernenden verstehen, was sie machen sollten und wo sie die benötigten Informationen her bekamen. Nachdem sich die Lernenden ausreichend mit den Repräsentationsformen auseinandergesetzt hatten erhielten sie die Aufgaben nacheinander auf einem gefalteten Blatt Papier. Dies sollten sie auffalten und die Aufgabenstellung zuerst alleine und leise lesen, bevor sie ihre Antworten in Partnerarbeit diskutieren sollten. Alle vier Aufgabenstellungen werden im Folgenden aufgeführt und im Anschluss einzeln erklärt:

- **Aufgabe 1:** Wenn du denkst, dass du die Abbildungen verstanden hast, diskutiere mit einem Partner die folgenden Punkte: Betrachte das Blatt mit dem Auge. Hier seht ihr drei unterschiedliche Abbildungen von der Netzhaut. Vergleicht die Abbildungen 3, 4 und 5 und erklärt, warum im Randbereich der Netzhaut, am gelben Fleck und am blinden Fleck nicht das Gleiche wahrgenommen werden kann.
- **Aufgabe 2:** Warum siehst du nachts alle Katzen grau?
- **Aufgabe 3:** Es wäre toll, wenn Menschen bei Nacht besser sehen könnten! Was müsste sich deiner Meinung nach beim Bau der Netzhaut ändern?
- **Aufgabe 4:** Mark hat letztes Folgendes erlebt: Es war Nacht, der Mond schien und er fuhr mit seinem Fahrrad einen unbeleuchteten Weg entlang. Leider war das Fahrradlicht kaputt und er hatte Schwierigkeiten den Weg zu erkennen. Anstatt direkt auf den Weg zu blicken, hat Mark dann auf den Rand des Weges geschaut und auf einmal konnte er den Weg besser erkennen. Mark konnte sich nicht erklären, warum er bei Dunkelheit den Weg besser erkennen konnte, wenn er nicht direkt darauf geschaut hat. Gib Mark hierfür eine mögliche Erklärung.

Um herauszufinden, wie die Schüler mit den Aufgaben umgehen und wo sie dabei Schwierigkeiten haben, wurden die Aufgaben bewusst unterschiedlich schwer konstruiert (siehe Schwierigkeitsgrade auf Seite 76 ff.). In Abhängigkeit zur jeweiligen Perspektive auf die Schwierigkeit von Aufgaben lassen sich die Aufgaben in die jeweiligen Kompetenzstufenmodelle der KMK und Mayer einordnen. In Tab. 18 erfolgt deren Einordnung nach den in der Biologiedidaktik momentan am Häufigsten verwendeten Modellen (siehe KMK 2005 und Mayer et al. 2013). Aus der Tabelle wird deutlich, dass sowohl der Schwierigkeitsgrad als auch der Komplexitätsgrad der Aufgaben im Interview steigt.

Tab. 18: Einordnung des Schwierigkeitsgrades der Aufgaben dieser Studie.

Aufgabe in dieser Studie	KMK-Anforderungsbereiche in Biologie, im Kompetenzbereich Fachwissen in Anlehnung an Leisen (siehe Tab. 12 auf Seite 81)	Kompetenzstufenmodell nach Mayer (siehe Tab. 13 auf Seite 83)
MC-Test zum Input	Stufe 1 - Kenntnisse zielgerichtet wiedergeben (Reproduktion)	Kompetenzstufe I
Aufgabe 1	Stufe 2 - Kenntnisse auswählen und anwenden (Anwendung)	Kompetenzstufe II (Mindeststandard)
Aufgabe 2	Stufe 2 - Kenntnisse auswählen und anwenden (Anwendung in bekanntem/unbekanntem Kontext)	Kompetenzstufe III (Regelstandard)
Aufgabe 3	Stufe 2 - Kenntnisse auswählen und anwenden (Anwendung in unbekanntem Kontext)	Kompetenzstufe IV (Regelstandard plus)
Aufgabe 4	Stufe 2 - Kenntnisse auswählen und anwenden (Anwendung in unbekanntem Kontext unter Berücksichtigung von mehr Inhalten als in Aufgabe 3)	Kompetenzstufe IV (Regelstandard plus)

Die Einordnung und Stufung der Aufgaben lässt eine grobe Einschätzung der Aufgaben zu. Für einen präziseren Vergleich muss jedoch jede Aufgabe genau betrachtet werden. Die ausführliche Darstellung der zur Lösung der Aufgaben notwendigen kognitiven Operationen sind ab Seite 235 aufgeführt.

5.8 Ablauf der Interviews

Nachdem die Testpersonen den Leistungsgruppen zugeteilt wurden, konnten sie in Zweiertteams (Dyaden) zum Interview eingeladen werden. Die Entscheidung für Dyaden geschah aus zwei Gründen: Erstens konnten die Schüler sich in den Partnerarbeitsphasen gegenseitig unterstützen, woraus auf indirektem Weg bei der Videoanalyse Rückschlüsse auf eventuelle Verstehenschwierigkeiten gezogen werden konnten. Zweitens führte die Dyade zu einer Entschärfung der für Schüler ungewohnten Interviewsituation.

Von Beginn an hat eine bewusste Trennung zwischen Lern- und Leistungssituation stattgefunden (Leisen 2006). Das resultierte und äußerte sich darin, dass in allen Interviews gezielt eine möglichst stress- und testfreie Atmosphäre geschaffen wurde. Das geschah über verbindliche Zusagen an die Schüler, die hier nur in der Kurzversion vorgestellt werden: *Keiner erfährt eure Namen, es gibt keine Noten, Fehler sind erlaubt, nichts von dem hier Gesagten verlässt diesen Raum. Wenn es Fragen gibt, dann fragt und schön dass Ihr da seid.* Alle Befragungen folgten einem halbstandardisierten Interviewleitfaden. Die Halbstandardisierung eignete sich besonders, weil gerade Vorstellungen und Verstehenschwierigkeiten von besonderem Forschungsinteresse waren.

An den Interviewtagen wurde die jeweilige Klasse morgens begrüßt und nochmals ausdrücklich betont, dass die bereits interviewten Schüler den noch Drankommenden nichts über den Inhalt verraten sollten. Danach wurden die betreffenden Lernenden von dem Interviewer nacheinander aus der Klasse geholt und in ein separates Zimmer

geführt. Hier wurden die Schüler erneut begrüßt. Nach der Klärung von Formalien und einiger Eisbrecherfragen zur Entspannung der Situation wurden die Schüler nach ihrem persönlichen Einverständnis gefragt und, ob sie sich in der Lage fühlen, bei dem Interview mitzumachen. Nach dem Einverständnis haben die Schülerdyaden abwechselnd einen der beiden Inputs bekommen. Die Schüler wurden aufgefordert in Einzelarbeit den Input (Seite 110 ff.) durchzulesen und zu sagen, wenn sie damit fertig sind. Im Anschluss daran wurden sie nach Verstehensschwierigkeiten gefragt und unbekannte Wörter erklärt. Im nächsten Schritt sollten die Schüler einen Multiple-Choice-Test (Seite 223) zum Inhalt des Input ausfüllen. Das Ausfüllen des Multiple-Choice-Tests geschah in Einzelarbeit. Die Schüler konnten hierbei jederzeit den Input zu Rate ziehen, sofern sie das wollten. Als Abschluss durften die Schüler ihre eigenen Antworten im Multiple-Choice-Test kontrollieren. Die Selbstkontrolle diente der Bestätigung, dass die Schüler hier nicht abgeprüft wurden, sondern selbst ihren Fortschritt kontrollieren konnten. In der nächsten Phase wurden den Schülern ergänzende Abbildungen (Seite 231 ff.) zum Input ausgehändigt. Genauso wie bei dem Multiple-Choice-Test sollten die Schüler den Arbeitsauftrag alleine durchführen. Hierbei mussten sie sich Abbildungen zum Auge betrachten und den besten Titel pro Abbildung aus einer Liste von möglichen Überschriften auswählen und ankreuzen. Danach wurden die gewählten Überschriften wieder gemeinsam kontrolliert und mögliche Unklarheiten bei den Abbildungen geklärt. Nach dem Bearbeiten der Abbildungen wurde den Schülern ein Modell vom Auge präsentiert (Seite 257). Hierbei sollten sie markante Strukturen aus dem Input und den Abbildungen im Modell wiedererkennen und zeigen. Im Anschluss an die Input-Phase bekamen die Lernenden nacheinander vier unterschiedliche Lernaufgaben vorgelegt (Seite 112 ff.). Die Schüler sollten versuchen, die Aufgaben in Partnerarbeit mündlich zu lösen. Das Ziel hierbei lag darin, den Umgang der Schüler mit Wissen bei unterschiedlich schwierigen Aufgaben zu untersuchen. Hierfür wurden im Vorfeld jeweils fünf Hilfen vorformuliert, wobei die letzte Hilfe der Lösung der Aufgabe entspricht (Seite 219 ff.). Die Hilfen lagen dem Interviewenden vor und kamen ausschließlich bei folgenden Bedingungen zum Einsatz:

- wenn Schüler von selbst Hilfe anforderten,
- wenn Schülerantworten in eine komplett falsche Richtung gingen,
- wenn Schüler an der Beantwortung eine Aufgabe zu scheitern drohten¹⁵.

Nach der Bearbeitung und Beantwortung aller Aufgaben wurde den Schülern am Ende des Interviews die Möglichkeit gegeben, die Aufgaben zu bewerten: Welche der vier

¹⁵ Dies geschah aus dem Grund, dass keine Schüler das Interview mit falschem Wissen verlassen durften oder das Gefühl des Scheiterns beim Interview zurück in den Unterricht bringen sollten.

Aufgaben war am einfachsten? Welche am schwersten? Welche war im Vergleich am interessantesten? Hierbei sollten sie nicht nur ihre Meinung äußern, sondern diese auch begründen. Im Anschluss wurde den Schülern für ihre Mitarbeit gedankt. Einen Überblick über den Ablauf der Interviews und den Einsatz der Materialien gibt Tab. 41 (Seite 220).

Das Dyaden-Dilemma

Obgleich gute Gründe für die Einladung von Schülerdyaden und deren Beforschung bestehen (siehe 5.8 auf Seite 113), besitzen Dyadeninterviews jedoch auch Nachteile (Bortz und Döring 2006, S. 246 ff.). So kam es schon bei der Erprobung der Interviews häufig vor, dass am Ende eines Interviewtages nur eine Person aus der Klasse übrig geblieben ist. In so einem Fall wurden die jeweiligen Schüler gefragt, ob sie es sich auch zutrauen würden, das Interview alleine durchzuführen.

Aus diesen Einzel-Interviews und den Gesprächen, davor und danach, zum Beispiel während die Schüler zurück in ihre Klasse gebracht wurden, konnten die folgenden Punkte abgeleitet werden, die das Dyaden-Dilemma treffend beschreiben (Bortz und Döring 2006, S. 84):

Das Dilemma bestand darin, dass Schüler in den Einzelinterviews entweder von vorneherein nicht mitmachen wollten¹⁶ oder kurz vor dem Interview sagten, dass sie es sich doch nicht zutrauen würden¹⁷. Dadurch reduzierte sich die Anzahl der Probanden. Zudem äußerte eine Schülerin, die alleine am Interview teilgenommen hatte, dass sie die Stille im Interview, wenn sie nachdenken musste manchmal als peinlich empfunden habe. Das hat sie zusätzlich unter Druck gesetzt, was dem Ziel aus der Planungsphase, eine möglichst stressfreie Atmosphäre während den Interviews zu schaffen, widersprach (Prüfer und Stiegler 2002).

Werden Schüler also einzeln eingeladen, so nehmen sie tendenziell unfreiwilliger an dem Interview teil, sind teilweise verschüchtert und generell eher zurückhaltender als bei einem Interview in Begleitung eines Partners. Bei einem Partnerinterview werden diese negativen Punkte zwar weitgehend entschärft, allerdings sinkt damit auch die Trennschärfe der Beobachtung. Diesen Sachverhalt verdeutlichen die folgenden Beispiele:

¹⁶ Ist während der Studie insgesamt drei Mal vorgefallen.

¹⁷ Ist während der Studie insgesamt zwei Mal vorgefallen.

- Schüler A und B lesen gemeinsam eine Aufgabe. Schüler A sagt sofort die Antwort, Schüler B sagt „So ist es!“. Hätte Schüler B die Antwort auch gewusst, wenn Schüler A diese nicht gesagt hätte?
- Zwei Schüler sitzen nebeneinander. Schüler A redet, Schüler B sagt die ganze Zeit nichts. Die Antwort von Schüler A ist unbefriedigend. Der Versuchsleiter möchte es genauer wissen und fragt nach. Schüler A weiß nicht weiter und Schüler B schweigt weiterhin. Der Versuchsleiter gibt Schüler A eine Hilfe. Schüler A liest die Hilfe vor, weiß die Antwort darauf allerdings nicht. Nun meldet sich Schüler B zu Wort. Er sagt zwar nicht die Antwort auf die Hilfe, gibt allerdings die Antwort auf die komplette Aufgabe. Hat Schüler B die Hilfe wirklich gebraucht oder benötigte er nur mehr Zeit? Hat er die Antwort gewusst oder hat ihn das Scheitern von Schüler A auf den richtigen Weg gebracht?

Aus diesen Beispielen wird ersichtlich, dass aus dem Dyaden-Setting einem Schüler eine Leistung oder ein Unvermögen unterstellt werden kann, wofür es keine eindeutigen Hinweise gibt. Hieraus ergibt sich eine mögliche Unschärfe in der Beobachtung beim Verhalten der Schüler.

In diesem Fall konnte dieser Unschärfe entgegen gewirkt werden, weil die Schüler mehrere Aufgaben über einen längeren Zeitraum gemeinsam bearbeiten mussten. Ein einfacher Vergleich des Verhaltens über all diese Aufgaben hinweg brachte schnell zum Vorschein, ob sich einer von den beiden bei der Lösung der Aufgaben zurückgehalten hat und nur die Antworten des Partners wiederholt oder sich produktiv an dem Interview beteiligt hat.

5.9 Methoden zur Auswertung der Daten

Entsprechend der Forschungsfragen (siehe Seite 3) und der Skalierung der erhobenen Messwerte, wurden für die Untersuchung der Daten unterschiedliche statistische Methoden verwendet. Diese werden als bekannt vorausgesetzt und im Folgenden in alphabetischer Reihenfolge aufgelistet. Neben einer kurzen Beschreibung, wozu die jeweilige Methode in diesem Projekt verwendet wurde, findet sich eine weiterführende Literaturangabe:

- ANOVA: Die zweifaktorielle Varianzanalyse ohne Messwiederholung wurde eingesetzt, um Haupt- bzw. Interaktionseffekte zu ermitteln. Die ANOVA wurde in Bezug auf das Multiple-Choice-Ergebnis der Schüler sowohl zwischen dem Input und den Leistungsgruppen als auch zwischen dem Input und dem Geschlecht der Schüler eingesetzt (siehe Seite 127 und 129 (Rasch et al. 2006b)).

- **Boxplot:** Bei allen Verfahren, die eine Normalverteilung der Daten voraussetzten, wurden diese per Boxplot zuerst grob gesichtet und dann genauer überprüft. Hierbei wurden vor allem Ausreißer identifiziert, die anschließend von den entsprechenden Berechnungen ausgeschlossen wurden (siehe Seite 139 (Williamson et al. 1989)). Ein Boxplot ermöglicht das direkte Ablesen von Daten aus der Abbildung (Abb. 21). Neben den Ausreißern (unerwartete Messwerte einer Messreihe) eines Datensatzes, beschreiben die beiden Antennen *Whisker* den kleinsten (Minimum) beziehungsweise größten Datenwert des Datensatzes (Maximum) und zeigen damit die Spannweite des Datensatzes. Der Beginn der Box (unteres Quartil) besagt, dass 25% aller Daten kleiner sind als dieser Wert und das Ende der Box (oberes Quartil) beschreibt, dass 75% aller Daten kleiner sind als dieser Wert. Im Vergleich zum arithmetischen Mittel eignet sich der Median für ordinal skalierte Daten. Hierbei teilt der Median einen Datensatz so in zwei Hälften, dass gleich viele Daten kleiner, beziehungsweise größer als der Median sind.

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 21: Beschreibung der Elemente eines Boxplots.

- **Chi-Quadrat-Test:** Um Häufigkeiten eines kategorialen Merkmals zwischen zwei Gruppen zu vergleichen, wurden Chi-Quadrat-Verfahren angewendet (siehe Seite 161 (Bortz und Döring 2006)).
- **Interrater-Reliabilität:** Um das Ausmaß der Übereinstimmung der Beurteilung von unterschiedlichen Beobachtern zu ermitteln, wurde eine Interrater-Reliabilität für die Einschätzung der Qualität der Schülerantworten berechnet (siehe Seite 122 (Krippendorff 2011)).
- **Korrelation:** Zur Überprüfung eines Zusammenhangs zweier intervallskalierter Daten wurden Korrelationen nach Pearson gerechnet (siehe Seite 169 (Bortz und Döring 2006)).
- **Levene-Test:** Zur Überprüfung der Voraussetzung der Gleichheit der Varianzen bei der zweifaktoriellen Varianzanalyse ohne Messwiederholung wurde der Levene-Test eingesetzt. Sofern keine Verletzung dieser Voraussetzung vorlag, wurden diese auch nicht im Text erwähnt (siehe Seite 127 und 129 (Fromm 2010)).
- **Lineare Regression:** Mit dieser Methode wurde die Auswirkung einer abhängigen Variable auf unabhängige Variablen überprüft, in diesem Fall die Auswirkung der Lernzeit auf das Multiple-Choice-Ergebnis (siehe Seite 136 (Kuckartz et al. 2010)).

- Reliabilitätsanalyse: Zur Ermittlung der inneren Konsistenz des Multiple-Choice-Tests wurde die Reliabilitätsanalyse mittels Cronbachs Alpha durchgeführt (siehe Seite 127 (Bortz und Döring 2006)).
- t-Test: Für den Vergleich der Mittelwerte zweier unabhängiger Gruppen wurden t-Test eingesetzt (siehe Seite 139 (Rasch et al. 2006a)).

Interrater Agreement zur Qualität der Schülerantworten

In diesem Abschnitt geht es um die Klärung der Übereinstimmung der Beurteiler bei der Einschätzung der Schülerleistungen in den Interviews. Hierfür wird zu Beginn eine Einführung in den theoretischen Hintergrund der Interrater-Reliabilitätsbestimmung gegeben und dargestellt, aus welchen Gründen die hier verwendeten Methoden zum Einsatz kamen. Im Anschluss wird die Umsetzung der Methode beschrieben und abschließend die Ergebnisse des Interrater Agreement genannt.

Theoretischer Hintergrund

Die Beurteilerübereinstimmung (interrater agreement) ermöglicht eine Aussage über das Ausmaß der Übereinstimmung unterschiedlicher Personen über ein Objekt in Bezug auf ein Merkmal (Cohen 1960). In diesem Fall entspricht die Beurteilerübereinstimmung der gleichen Einschätzung der Qualität der Schülerantworten anhand von Kriterien. Zur Bestimmung der Übereinstimmung und damit der Qualität der erhobenen Messdaten können unterschiedliche Methoden eingesetzt werden. Die Wahl der Methode hängt dabei hauptsächlich von der Skalierung der erhobenen Daten ab (Wirtz 2002). Neben der Skalierung ist es ferner von Relevanz, ob die Korrelation der Werte zwischen zwei oder mehr Ratern von Interesse ist, oder ob es um die tatsächliche Übereinstimmung der Einschätzungen von zwei oder mehr Beurteilern geht. Die hierfür in Frage kommenden Analysemethoden sind in Tab. 19 dargestellt.

Tab. 19: Analysemethoden zur Beurteilerübereinstimmung in Abhängigkeit zur Skalierung der Daten (in Anlehnung an Dorsch et al. 2014).

Skalierung der Daten	Analysemethode
dichotom	Cohens Kappa*, Fleiss' Kappa**
nominal	Cohens Kappa*
ordinal	gewichtetes Kappa*, Spearmans Rangkorrelationskoeffizient*, Kendalls Tau*, Kendalls Konkordanzkoeffizient**, Krippendorffs Alpha**
intervall	Intra-Klassen-Korrelation**

* Für zwei Beobachter

** Für mehr als zwei Beobachter

Der Grund für die Berechnung eines Interrater Agreements dient aber nicht nur der statistischen Überprüfung, ob es sich bei der Einschätzung der Bewerter um Zufall handelt, sondern vor allem, dass die Coder bei ihren Bewertungen nicht zu weit auseinanderliegen und ganz besonders, dass das Ergebnis vergleichbar und replizierbar ist (Krippendorff 2004, S. 237). Daher sollten für die Einschätzung der Daten über die statistischen Ergebnisse hinaus noch weitere Angaben über das Interrater Agreement gemacht werden, die die Durchführung nachvollziehbarer machen und die Interpretation erleichtern.

Hierfür wurden bei einer Vergleichsanalyse von 96 Studien acht Kriterien identifiziert, die den Vergleich und die Replizierbarkeit von Studien erheblich erhöhen, sofern sie in zukünftigen Studien auch genannt werden (Bayerl und Paul 2011). Diese Kriterien werden im Folgenden aufgeführt und erklärt:

- 1) **Domäne:** Je nach Arbeitsgebiet können sich Beurteilungen gravierend voneinander unterscheiden. Das gilt sowohl für die Form der Daten (Transkripte, Audiomitschnitte, Videosequenzen), als auch für das Fach- und Wissensgebiet, in dem sich die zu bewertenden Daten bewegen.
- 2) **Anzahl der Beobachter:** Bei der Beschreibung des Interrater Agreements ist die Anzahl der Beobachter nicht nur für die Auswahl der Analysemethode notwendig. Sie spielt schon bei der Planung der Studie eine Rolle, denn mit zunehmender Anzahl an Beurteilern sinkt deren Übereinstimmung.
- 3) **Training:** Die Schulung der Bewerter verbessert deren Übereinstimmung. Das ist vor allem bei intensivem Training der Fall.
- 4) **Zweck der Beurteilung:** Die Konsequenz der Beurteilungen geben die Mindest-Qualität der Übereinstimmung vor. Bei klinischen Studien müssen andere Cut-Offs bei der Bestimmung der Beurteilerübereinstimmung gelten als bei Reliabilitätsanalysen in Politik- oder Sozialwissenschaften.
- 5) **Expertise:** Je nach Erfahrungs- und/oder Fachwissen können sich die Beurteilungen zwischen Experten und Novizen stark voneinander unterscheiden.

Hierbei erreichen homogene Gruppen aus Experten oder Novizen in der Regel höhere Übereinstimmungen als heterogene Beurteilergruppen.

- 6) **Sprache:** Besonders in Bereichen der Kommunikation sollte drauf geachtet werden, dass die Sprache, in der die Daten vorliegen, der Muttersprache der Beurteiler entspricht oder diese auf einem sehr hohen Niveau beherrschen, damit sie auch zwischen den Zeilen lesen können.
- 7) **Anzahl der Kategorien:** Mit zunehmender Anzahl der Bewertungskriterien sinkt die Übereinstimmungsrate.
- 8) **Analysemethode:** Die Art der Berechnung des Interrater Agreements hat systematischen Einfluss auf das Ergebnis. Mehrheits- und paarweise Vergleichsmethoden führen dabei zu höheren Werten als Konsensmethoden oder Methoden zur Beschreibungen der absoluten Übereinstimmung.

Methodische Umsetzung

Für die Ermittlung der Interrater-Reliabilität in dieser Studie wurden 25 der 56 Interviews per Zufallsgenerator¹⁸ ausgewählt. Exemplarisch wurde die Interraterreliabilität für Aufgabe 4 untersucht. Den Bewertern wurde hierfür ein Video (Dauer: 19 Minuten und 12 Sekunden) mit dem folgenden Auftrag vorgelegt:

- 1) Lesen Sie zuerst diese Anleitung und die Kodierungsregeln aufmerksam durch.
- 2) Schauen Sie sich anschließend das Video in voller Länge an, damit Sie ein Gefühl für die Qualität der Schülerantworten entwickeln.
- 3) Gehen Sie im Anschluss das Video erneut durch und notieren Sie Ihre Einschätzung der Schülerantwort entsprechend der folgenden Bewertungskriterien:
 - a) Eine Antwort ist als *gut* zu markieren, wenn die Person eigenständig eine Antwort sagt, diese in ganzen Sätzen und zusammenhängend formuliert und die richtigen Fachwörtern verwendet.
 - b) Eine Antwort ist als *schlecht* zu markieren, wenn die Person zwar eine Lösung sagt, diese aber nicht in ganzen Sätzen oder ungenau formuliert ist, die Person falsche Begriffe benutzt oder nur einen Teil der Antwort gibt.
 - c) Eine Schülerantwort ist der Kategorie *gelingt nicht* zuzuordnen, wenn die Antwort falsch ist, wenn die Person nichts sagt, den Versuchsleiter eigenständig nach der Lösung bittet, bei der Aufgabe aufgibt oder der Versuchsleiter die Lösung nach angemessener Zeit und Hilfe geben muss.

Bevor auf die Ergebnisse der Beurteilerübereinstimmung eingegangen wird, werden in Tab. 20 relevante Informationen in Bezug auf die Kriterien von Bayerl und Paul genannt.

¹⁸ Online Zufallsgenerator: <http://www.zufallsgenerator.net> (zuletzt geprüft am 27.03.2014).

Tab. 20: Umsetzung der Kriterien nach Bayerl (2011) zur Einschätzung der Beurteilerübereinstimmung in dieser Studie (Interrater Agreement).

Domäne: Die Daten liegen als Video vor. Die zu bewertenden Antworten der Schüler beziehen sich ausschließlich auf biologische Fachinhalte der Sekundarstufe I in Baden-Württemberg.

Anzahl der Beobachter: 3

Training: Die Bewerber erhielten kein Training. Die einzige Vorbereitung der Bewerber lag in der Durchsicht aller Interviews, ohne diese zu bewerten (Dauer ca. 19 min). Die Gesamtdauer der Bewertung belief sich auf ca. 45 Minuten.

Zweck der Beurteilung: Bei der Beurteilerübereinstimmung ging es in dieser Studie um die Überprüfung der Resultate durch die verwendeten Kodierungsregeln. Dies geschah durch die Replizierung der Ergebnisse durch andere Personen.

Expertise: Bei den drei Bewertern handelt es sich um einen Experten und zwei Novizen*.

Sprache: Sowohl das Datenmaterial als auch die Muttersprache der Bewerber war auf Deutsch.

Anzahl der Kategorien: Für die Bewertung der Schülerantworten lagen 3 Kategorien vor.

Analysemethode: Entsprechend der statistischen Analysemöglichkeiten von ordinalskalierten Daten und mehr als zwei Ratern wurden sowohl Kendalls Konkordanz Koeffizient als auch Krippendorffs Alpha¹⁹ berechnet.

** Bei den Novizen handelt es sich um Personen, die die Aufgaben in den Videos lösen konnten, aber wenig bis keine Erfahrung im Umgang mit Schülern haben (keine Lehrer oder Erzieher) und der deutschen Sprache mächtig sind. Bei den Beurteilern handelt es sich um eine Frau und einen Mann, beide studiert (Betriebswirtschaftslehre und Rechtswissenschaft) und in der Altersgruppe zwischen 25 und 30 Jahren.*

Ergebnis zur Beurteilerübereinstimmung

Bei der statistischen Untersuchung sind sowohl bei der Berechnung des Kendalls Tau sowie des Krippendorffs Alpha (0.746) gute bis akzeptable Werte für die Übereinstimmung der Bewerber herausgekommen (siehe Tab. 21). Auch in dieser Studie unterschieden sich die Bewertungen von Experten und Novizen. Bei einem Vergleich der Ergebnisse von Novize 1 mit Novize 2 korrelierten die Ergebnisse stärker miteinander als die Bewertung des Experten im Vergleich zu den Novizen, was sich mit Ergebnissen von Bayerl und Paul deckt (Bayerl und Paul 2011, S. 715). Durch die Überprüfung mit zwei Analysemethoden konnte gezeigt werden, dass unter Verwendung der eingesetzten Kodierungsregeln auch ungeschulte Nicht-Experten zu sehr ähnlichen Ergebnissen kommen.

¹⁹ Die Berechnung des Krippendorffs Alpha ist in der Standardversion von SPSS 21 nicht möglich. Zu dessen Berechnung gibt es zwei Möglichkeiten: (1) die Installation des Macros KALPHA (Hayes und Krippendorff 2007 downloadlink: <http://www.afhayes.com/spss-sas-and-mplus-macros-and-code.html>) oder (2) dessen Berechnung online (Freelon 2013 siehe <http://dfreelon.org/utis/recalfront/>). Die Werte dieser Studie wurden mit dem Onlineverfahren ermittelt (beide hier aufgeführte Links wurden zuletzt geprüft am 04.05.2014).

Tab. 21: Nichtparametrische Korrelationen nach Kendalls Tau (b) zur Beurteilerübereinstimmung (signifikante Werte sind fett gedruckt).

		Experte	Novize 1	Novize 2
Experte	Korrelationskoeffizient	1,000	,623**	,685**
	Sig. (2-seitig)	.	,000	,000
	N	48	48	48
Novize 1	Korrelationskoeffizient	,623**	1,000	,795**
	Sig. (2-seitig)	,000	.	,000
	N	48	48	48
Novize 2	Korrelationskoeffizient	,685**	,795**	1,000
	Sig. (2-seitig)	,000	,000	.
	N	48	48	48

** Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

6 EMPIRISCHE UNTERSUCHUNG

6.1 Einleitung

In den folgenden Abschnitten werden die Ergebnisse dieser Studie aufgeführt, die auf den Entscheidungen und Materialien aus der Zusammenarbeit mit Lehrkräften beruhen (siehe Seite 97 ff.). Die Reihenfolge der nachfolgenden Unterkapitel entspricht den Forschungsfragen (siehe Tab. 22). Jedes Unterkapitel ist gegliedert in Fragestellung, methodisches Vorgehen und Ergebnis.

Bei den Ergebnissen werden sowohl alle signifikanten als auch nicht signifikanten Ergebnisse aufgeführt. Gegen Ende dieses Kapitels werden die Forschungsfragen durch weitere Fragen ergänzt, die während dem Forschungsprozess entstanden sind.

Tab. 22: Zusammenstellung der hier bearbeiteten Forschungsfragen.

Forschungsfragen	Seite
Welche Ansätze innerer Differenzierung eignen sich aus fachdidaktischer Sicht für die Vermittlung biologischer Inhalte? Welche dieser Ansätze nehmen Lehrerinnen und Lehrer aufgrund ihres Erfahrungswissens an?	97
Hat eine differenzierende Darstellungsform annähernd identischer Lehrinhalte über Experten-Concept-Maps oder Text Auswirkungen auf das Wissen und Können bei Lernenden mit unterschiedlichen Leistungsniveaus?	126
Treten bei den eingesetzten Differenzierungsmaßnahmen geschlechtsspezifische Unterschiede auf?	129
Welchen Einfluss hat die Lernzeit auf das Ergebnis bei Schülern?	132
Wie gestaltet sich der Lernprozess der Schüler mit den differenzierten Darstellungsformen und der Lernaufgaben und wo treten dabei Schwierigkeiten auf?	144
Fragestellungen, die über die primären Forschungsfragen hinausgehen	154
- Wie schätzen Schüler den Schwierigkeitsgrad der Aufgaben ein und welche Aufgabe finden Schüler am interessantesten?	154
- Wie fällt die Qualität der Schülerantworten bei unterschiedlichen Schülergruppen aus?	161
- Bei welchen Abbildungen hatten die Schüler die meisten Probleme?	167

6.2 Erkenntnisse aus der Erprobung der Materialien

6.2.1 Schüler verfolgen individuelle Lernwege

Bei der Erprobung der Interviews hat sich gezeigt, dass die Bearbeitung der Hilfen in strikter Reihenfolge (1 bis 5) in dem Setting einer Interviewsituation unvorteilhaft war. Dies zeigte sich daran, dass Schüler durch die Impulse der Hilfen auf andere Dinge hingewiesen wurden, als solche, über die sie gerade diskutierten. Anders gesagt, die Schüler sind ganz unterschiedlich an die Lösung der Aufgaben herangegangen, als es geplant war. Die Einhaltung der Chronologie der Hilfen bedeutete für manche Schüler einen Gedankensprung zu machen, ihre eigenen Gedanken zu übergehen oder gar in ihrer Argumentation einen Rückschritt zu machen. Daher wurde von der regulären Verfahrensweise abgewichen und die Hilfen nicht hintereinander sondern angepasst an den Lernprozess gegeben.

6.2.2 Das Augenmodell fördert Verstehen und Kommunikation

Nach der Erprobung der Materialien wurden die Schüler gebeten, ihre Meinung zu dem Interview zu äußern. Dabei kam heraus, dass Schüler bei manchen der Aufgaben die Antwort schon gewusst hatten, bevor sie die Musterlösung erhalten haben, die Antwort allerdings nicht gesagt haben, weil sie große Schwierigkeiten hatten, ihre Antwort zu formulieren. Das war ein Problem, denn wenn es schon bei der Erprobung der Interviews dazu kommt, dass Schüler die Antwort zwar wissen, sie aber nicht sagen, weil ihnen die passenden Worte fehlen, würde dies das Ergebnis der Studie verfälschen. Aus diesem Grund wurde in den folgenden Interviews ein Modell vom menschlichen Auge hinzugenommen.

Das Modell sollte es den Schülern erleichtern, sich den Aufbau des Auges räumlich vorzustellen und anhand des Modells Sachverhalte zu erklären, wofür ihnen sonst die Worte fehlen (zum Beispiel mit dem Kugelschreiber den Lichteinfall ins Auge am Modell zeigen). Mit der nachträglichen Hinzunahme des Augenmodells in den Ablauf der Interviews sind nicht nur weitere kognitive Anforderungen an die Schüler verbunden (siehe Seite 242), sondern es trat auch eine weitere Besonderheit auf, die sich erst während der Datenerhebung gezeigt hat: Viele Schüler verwendeten das Modell nicht. Sie fassten es weder an, noch nahmen sie es auseinander, obwohl sie explizit dazu aufgefordert wurden. Dies kann an der ungewohnten Situation gelegen haben oder daran, dass die meisten Lernenden im Umgang mit Modellen bis zu diesem Zeitpunkt nicht vertraut waren.

6.3 Experten-Concept-Maps im Vergleich zu Text

6.3.1 Alle Schüler profitieren von Experten-Concept-Maps

Die Experten-Concept-Map ist in Kooperation mit Lehrern im Bereich der Reproduktion entwickelt worden. Im Gegensatz zum Fachtext handelt es sich bei der Experten-Concept-Map um eine abweichende Form des Inputs, der das Hauptinteresse der Studie galt.

Frage: Hat eine differierende Darstellungsform annähernd identischer Lehrinhalte Auswirkungen auf das Wissen von Schülern?

Beschreibung der Vorgehensweise: Um der Frage nachzugehen, wurden Schüler entweder mit einem Text oder einer Experten-Concept-Map konfrontiert. Um die Auswirkung des Inputs auf die Schüler zu untersuchen, mussten sie im Anschluss einen Multiple-Choice-Test ausfüllen (siehe Seite 223).

Generelle Aspekte: Wegen der Streichung von vier Ausreißern reduzierte sich die Populationsgröße für die folgenden Berechnungen von 112 auf 108. Die Verteilung der Geschlechter war annähernd gleich zwischen beiden Input-Gruppen: 23 Jungen erhielten einen Text und 25 Jungen eine Experten-Concept-Map, bei den Mädchen erhielten 31 einen Text und 29 eine Experten-Concept-Map. Ungeachtet der Gruppenzugehörigkeit machte kein Schüler 0 Punkte. Das tatsächlich erzielte Multiple-Choice-Ergebnis reichte von 15 (Min) bis 28 (Max) Punkte, mit einem Mittelwert von 23,19 und einem Median von 24 Punkten. Die Testreliabilität zeigte bei den Rohdaten eine gute interne Konsistenz mit einem Cronbachs Alpha von .74. Die relevanten Mittelwerte und Standardabweichungen sind in Tab. 23 aufgeführt.

Ergebnis: Experten-Concept-Maps zeigten einen positiven Effekt auf das Wissen von Schülern. Hierfür wurde eine zweifaktorielle Varianzanalyse ohne Messwiederholung gerechnet (two-way ANOVA). Dabei wurden die Varianzen der Multiple-Choice-Ergebnisse (abhängige Variable) in Abhängigkeit zum Input (Text oder Experten-Concept-Map) und zu den Leistungsgruppen (leistungstärkere und leistungsschwächere Schüler) miteinander verglichen. Hierbei zeigte sich ein Haupteffekt vom Input auf das Multiple-Choice-Ergebnis (siehe Tab. 24): alle Schüler erzielten mit der Experten-Concept-Map signifikant bessere Ergebnisse als mit dem Text ($p = .018$, $\eta^2 = .052$). Die Zugehörigkeit zu einer der beiden Leistungsgruppen zeigte einen marginal signifikanten Interaktionseffekt zwischen dem Input und der Leistungsgruppe mit einer entsprechend schwachen Effektstärke ($\eta^2 = .021$).

6.3.2 Leistungsschwächere Schüler profitieren marginal

Um dieser Effektstärke nachzugehen, wurde im Anschluss eine einfaktorische Varianzanalyse (ANOVA) gerechnet, bei der die Multiple-Choice-Ergebnisse der vier Untergruppen miteinander verglichen wurden. Das Ergebnis dieser ANOVA zeigte einen marginal signifikanten, förderlichen Effekt ($p = .058$) von Experten-Concept-Maps auf das Multiple-Choice-Ergebnis von leistungsschwächeren Schülern im Vergleich zu leistungsschwächeren Schülern mit Text (siehe Tab. 25).

Tab. 23: Deskriptive Daten zum Multiple-Choice-Ergebnis bei leistungsstärkeren und leistungsschwächeren Schülern in Abhängigkeit vom Input.

Input	Leistungsgruppen	<i>M</i>	<i>SD</i>	<i>N</i>
Text	Leistungsstärkere	22.83	2.705	30
	Leistungsschwächere	22.17	2.914	24
	Gesamt	22.54	2.793	54
Experten-Concept-Map	Leistungsstärkere	23.34	3.243	29
	Leistungsschwächere	24.40	2.958	25
	Gesamt	23.83	3.131	54
Gesamt	Leistungsstärkere	23.08	2.967	59
	Leistungsschwächere	23.31	3.117	49
	Gesamt	23.19	3.024	108

Tab. 24: Das Ergebnis der zweifaktoriellen Varianzanalyse ohne Messwiederholung (Input und Leistungsgruppe) zeigte einen signifikanten Effekt von Experten-Concept-Maps auf das Multiple-Choice-Ergebnis aller Schüler (signifikante Ergebnisse sind fett gedruckt, interessante Ergebnisse sind fett und kursiv gedruckt).

	<i>df</i>	<i>F</i>	<i>p</i>	η^2
Input	1	5.747	.018	.052
Leistungsgruppe	1	.115	.735	.001
Input * Leistungsgruppe	1	2.262	.136	.021
Fehler	104			
Gesamt	108			
Korrigierte Gesamtvarianz	107			

Tab. 25: Das Ergebnis der einfaktoriellen ANOVA zeigte einen marginal signifikanten Effekt von ECMs (ECM = Experten-Concept-Map) auf leistungsschwächer Schüler. Entsprechende Ergebnisse sind fett gedruckt.

		<i>Mittlere Differenz</i>	<i>SE</i>	<i>p</i>	<i>95%-CI von - bis</i>	
Leistungsstärkere mit Text	Leistungsstärkere mit ECM	-.511	.771	1.000	-2.59	1.56
	Leistungsschwächere mit Text	.667	.811	1.000	-1.51	2.85
	Leistungsschwächere mit ECM	-1.567	.802	.321	-3.72	.59
Leistungsstärkere mit ECM	Leistungsstärkere mit Text	.511	.771	1.000	-1.56	2.59
	Leistungsschwächere mit Text	1.178	.817	.914	-1.02	3.38
	Leistungsschwächere mit ECM	-1.055	.808	1.000	-3.23	1.12
Leistungsschwächere mit Text	Leistungsstärkere mit Text	-.667	.811	1.000	-2.85	1.51
	Leistungsstärkere mit ECM	-1.178	.817	.914	-3.38	1.02
	Leistungsschwächere mit ECM	-2.233	.846	.058	-4.51	.04
Leistungsschwächere mit ECM	Leistungsstärkere mit Text	1.567	.802	.321	-.59	3.72
	Leistungsstärkere mit ECM	1.055	.808	1.000	-1.12	3.23
	Leistungsschwächere mit Text	2.233	.846	.058	-.04	4.51

6.4 Genderaspekte

6.4.1 Mädchen haben bessere Noten, aber keine besseren Ergebnisse

Während der Suche nach Unterschieden zwischen leistungsstärkeren und leistungsschwächeren Schülern traten Kontroversen auf, die das Geschlecht der Schüler, deren Leistungsgruppe und deren Ergebnisse im Multiple-Choice-Test betrafen. Zur Erinnerung: Die Schüler wurden ausschließlich anhand ihrer Biologienote in Leistungsgruppen aufgeteilt. Aufgrund dieser Aufteilung setzte sich die Gruppe der leistungsstärkeren Schüler vornehmlich aus Mädchen zusammen (39 ♀ vs. 20 ♂), wohingegen in der Gruppe der leistungsschwächeren Schüler die Jungen dominierten (28 ♂ vs. 21 ♀). Basierend auf der Einschätzung der Schüler über ihre Note wären bei den Mädchen im Vergleich zu den Jungen höhere Punktzahlen zu erwarten gewesen. Dies trat allerdings nicht ein. Die von den Mädchen tatsächlich erzielten Ergebnisse bewegten sich in eine unerwartete Richtung.

6.4.2 Beide Geschlechter profitieren von Experten-Concept-Maps

In Bezug auf die hier verwendete Differenzierungsmaßnahme ergab sich zuerst folgende Frage:

Fragestellung: Hat der Input unterschiedliche Auswirkungen auf das Ergebnis von Mädchen und Jungen?

Methodisches Vorgehen: Um Effekte des jeweiligen Inputs und des Geschlechts auf das Multiple-Choice-Ergebnis zu finden, wurde eine weitere zweifaktorielle ANOVA gerechnet. Hierbei diente das Multiple-Choice-Ergebnis als abhängige Variable und die beiden Faktoren (Geschlecht und Input) als unabhängige Variable.

Ergebnis: Die zweifaktorielle ANOVA ohne Messwiederholung zeigte einen statistisch signifikanten Haupteffekt: Beide Geschlechter profitieren gleichermaßen von Experten-Concept-Maps ($p = .033$, $\eta^2 = .043$). Das deckt sich mit anderen Ergebnissen in dieser Arbeit (siehe Tab. 24 auf Seite 128). Es zeigte sich allerdings kein Haupteffekt zwischen Mädchen und Jungen und kein Interaktionseffekt zwischen dem verwendeten Input und dem Geschlecht der Schüler (siehe Tab. 26).

Tab. 26: Das Ergebnis der zweifaktoriellen ANOVA zeigt einen Haupteffekt des Inputs. Es konnte weder ein Haupteffekt durch das Geschlecht, noch ein Interaktionseffekt zwischen Geschlecht und Input gefunden werden. Dies bedeutet, dass sich der Input auf beide Geschlechter positiv auswirkt. Signifikante Werte sind fett gedruckt.

	<i>df</i>	<i>F</i>	<i>p</i>	η^2
Geschlecht	1	.595	.442	.006
Input	1	4.661	.033	.043
Geschlecht * Input	1	.242	.623	.002
Fehler	104			
Gesamt	108			
Korrigierte Gesamtvarianz	107			

6.4.3 Das Geschlecht hat Einfluss auf Schulnoten

Der einzige resultierende Unterschied zwischen Mädchen und Jungen, der die Diskrepanz zwischen dem erwarteten und dem tatsächlichen Ergebnis erklärt, ist die Zuteilung der Schüler zu den Leistungsgruppen über deren Biologienote. Hieraus ergab sich zwangsläufig folgende Frage:

Fragestellung: Gibt es einen Unterschied zwischen den Geschlechtern in Bezug auf deren Biologienote?

Methodisches Vorgehen: Um den Unterschied zwischen zwei unabhängigen Populationen in Bezug auf ein Merkmal zu ermitteln wurden t-Tests gerechnet. Da nicht nur die Note der Schüler in Biologie in dieser Studie erhoben wurde, sondern auch die Deutsch- und Mathematiknote, wurden diese mit berechnet.

Ergebnis: Die Ergebnisse der t-Tests zeigten, dass statistisch signifikante Unterschiede zwischen Mädchen und Jungen in Bezug auf deren Noten in den Fächern Biologie und Deutsch bestehen. Es konnte kein bedeutender Notenunterschied in Mathematik gefunden werden (siehe Tab. 27).

Tab. 27: Vergleich der Schulnoten in Bezug auf das Geschlecht der Schüler. Sowohl in Biologie als auch in Deutsch bestehen signifikante Notenunterschiede zwischen den Geschlechtern, außer in Mathematik. Signifikante Werte (p) sind fett gedruckt, sowie deren Effektstärken (d) aufgeführt.

Schulfach	Geschlecht	N	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>p</i>	<i>d</i>
Biologie	♂	48	2.73	.609	1.30	3.75	.013	.49
	♀	60	2.43	.612	1.00	4.00		
Deutsch	♂	48	3.14	.788	1.50	4.50	< .001	.84
	♀	60	2.57	.576	1.50	4.25		
Mathe	♂	48	2.90	.828	1.30	4.50	.17	.26
	♀	60	2.70	.717	1.25	4.50		

6.5 Ergebnisse zum Faktor Zeit

Bei den Interviews kam es in Bezug auf die Zeit immer wieder zu Auffälligkeiten wie zum Beispiel bei der zeitlich sehr unterschiedlichen Auseinandersetzung mit dem Input und dem Ausfüllen des Multiple-Choice-Tests. Dies führte schon während der Datenerhebungsphase zu dem Verdacht, dass Zeit eine größere Rolle spielt als anfänglich gedacht. Dieser Verdacht erhärtete sich bei der deskriptiven Beschreibung der Daten und führte zu weiteren Forschungsfragen, auf die im Folgenden eingegangen wird.

6.5.1 Der Zeitbedarf ist individuell verschieden

Deskriptive Beschreibung der Zeit: Die Gesamtdauer aller Interviews von der Begrüßung bis zur Verabschiedung beträgt ca. 40 Stunden. Die durchschnittliche Dauer der Interviews lag bei 42 Minuten. Das kürzeste Interview dauerte 23 Minuten und das längste 65 Minuten. Dies bedeutet, dass eine Dyade für das gleiche Interview fast 3-mal soviel Zeit benötigte wie eine andere. Dieser Unterschied führte zu der ersten post-hoc Frage in Bezug auf die Zeit.

Fragestellung: Wofür benötigen Schüler im Interview mehr Zeit als andere?

Methodisches Vorgehen: Zur effektiveren Bearbeitung in der Videoanalyse wurden die Interviews in Phasen unterteilt. Bei diesen Phasen handelt es sich um die Begrüßung mit Formalitäten und Warm-up, die individuelle Bearbeitung des Inputs mit anschließendem schriftlichem Multiple-Choice-Test, die mündliche Bearbeitung von vier Lernaufgaben mit gestuften Hilfen in Partnerarbeit gefolgt von der Rückmeldung der Schüler zu den Aufgaben mit Danksagung und Verabschiedung. Bei einem groben Vergleich aller Phasen in den Interviews hat sich gezeigt, dass sich die Zeiten besonders bei der Bearbeitung der Lernaufgaben voneinander unterscheiden, was zu einer genaueren Betrachtung der Zeit dieser Phase der Interviews führte. Für die Kodierung der Zeit, die die Schüler pro Aufgabe benötigten, galt bei allen Videos das Aufnehmen, Öffnen und Lesen des Aufgabenzettels als Beginn der Zeitmessung und die Lösung als Ende. Die daraus resultierenden vier Abschnitte wurden pro Dyade in Sekunden kodiert und aufgelistet.

Ergebnis: Ein Vergleich der Mittelwerte (siehe Tab. 28) zeigte, dass die Schüler im Durchschnitt am meisten Zeit für Aufgabe 1 benötigen und für Aufgabe 2 am wenigsten. Der letztendlich viel interessantere Punkt bei der Aufschlüsselung der Zeit war allerdings nicht der Unterschied zwischen den Aufgaben, sondern der Unterschied zwischen den Dyaden pro Aufgabe. Diese Werte machen deutlich, welche enormen Unterschiede zwischen den Schülern in Bezug auf deren benötigte Zeit zum Lösen der Aufgaben bestanden.

Tab. 28: Zeitangaben zu der Dauer der Aufgaben 1 bis 4 in Sekunden (gerundet).

	Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgabe 4
Min	153	26	45	62
Max	868	293	341	763
Mittelwert	373	119	154	320
Median	320	113	134	299
Standardabweichung	167	61	67	149

6.5.2 Leistungsschwächere Schüler benötigen mehr Unterstützung

Diese Unterschiede zwischen den Dyaden führten dazu, sich mit der Komponente Zeit in der Videoanalyse tiefergehend zu beschäftigen. Abgesehen von den qualitativen Aussagen, die sich über das Lösen der Aufgaben treffen lassen (siehe Seite 126 ff.), ermöglicht die Videoanalyse eine weitere quantitative Aussage. Diese bezieht sich auf die reine Sprechzeit der beteiligten Personen und ermöglicht die Beantwortung der folgenden Frage:

Fragestellung: Wer spricht im Interview wie lange?

Methodisches Vorgehen: Hierfür wurden pro Video nur die Stellen markiert, an denen die Schüler oder der Versuchsleiter tatsächlich sprechen, also keine Denk- und Sprechpausen machen oder Füllwörter benutzen. Dies wurde exemplarisch für Aufgabe 4 durchgeführt.

Ergebnis: Das Ergebnis für die Sprechzeit der Schüler (siehe Tab. 29) zeigt das selbe Bild wie die Dauer für die Aufgaben (siehe Tab. 28): die Schüler unterschieden sich in Bezug auf den Zeitbedarf enorm. Bei beiden Tabellen wird vor allem bei dem Vergleich der minimalen mit den maximalen Werten der Unterschied zwischen den Schülern sehr deutlich. Dieser Unterschied wird bei der Betrachtung der Sprechzeit noch größer. Hierzu muss gesagt werden, dass die Interviews in Partnerarbeit durchgeführt wurden. Die teilweise sehr geringen minimalen Werte von unter einer Sekunde Sprechzeit entstanden dadurch, dass ein Schüler die Antwort sagte und der andere der Lösung nur zustimmte.

Tab. 29: Sprechzeit des Versuchsleiters in den Interviews (n=56) und die Sprechzeiten der Schüler (n=108) bei Aufgabe 4 in Sekunden.

	Sprechzeit des Versuchsleiters	Sprechzeiten der Lernenden
Min	4,2	0,7
Max	232,6	319,2
Mittelwert	115,6	56,6
Median	112,6	40,6
Standardabweichung	60,5	50,7

Es wurde eine Auswahl an besonders langsamen und besonders schnellen Schülern in Bezug auf die Dauer der Aufgaben und Sprechzeit erneut gesichtet. Dies geschah, um nach bisher unbedachten Besonderheiten zu suchen. Hierbei wurde ein Punkt besonders deutlich: Bei dem Vergleich der Sprechzeit von leistungsstärkeren und leistungsschwächere Schülern kam heraus, dass leistungsstärkere Schüler im Durchschnitt etwas mehr im Interview sprechen und der Versuchsleiter weniger. Im Gegensatz dazu sprechen bei annähernd identischer Aufgabendauer leistungsschwächere Schüler im Durchschnitt etwas weniger, dafür aber der Versuchsleiter deutlich mehr (siehe Tab. 30, Abb. 22 und Abb. 23).

Tab. 30: Vergleich der Sprechzeiten des Versuchsleiters und der Schüler in Sekunden (Aufgabe 4).

		Sprechzeit des Versuchsleiters	Sprechzeit der Lernenden
Leistungsstärkere Schüler	N (Anzahl der Interviews)	31	59
	Mittelwert	108.2	52.5
	Median	97.8	38.8
	Standardabweichung	59.5	38.7
	Minimum	4.2	.8
	Maximum	232.6	143.2
Leistungsschwächere Schüler	N (Anzahl der Interviews)	24	49
	Mittelwert	124.2	45.2
	Median	137.2	39.7
	Standardabweichung	62.9	33.4
	Minimum	18.0	.7
	Maximum	223.6	142.3

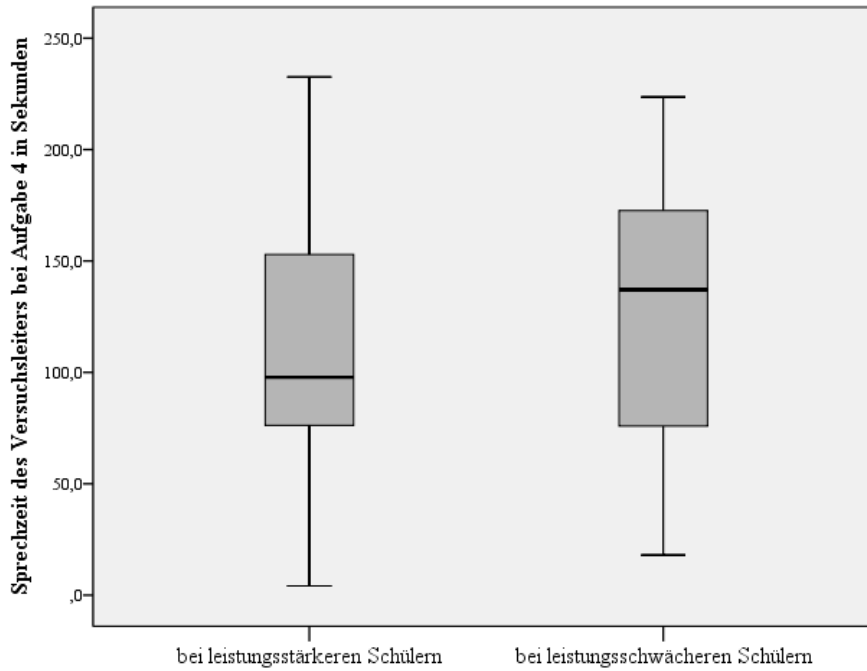


Abb. 22: Sprechzeit des Versuchsleiters bei leistungsstärkeren Schülern (n=31) im Vergleich zu leistungsschwächeren Schülern (n=24) als Boxplot (siehe Abb. 21 auf Seite 117).

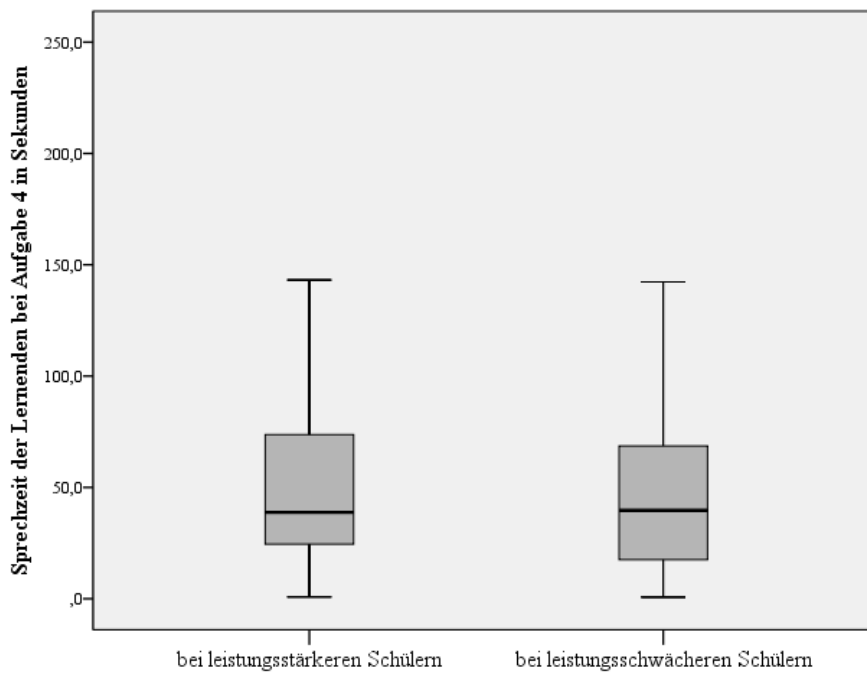


Abb. 23: Sprechzeiten der Lernenden bei leistungsstärkeren Schülern (n=59) im Vergleich zu leistungsschwächeren Schülern (n=49) als Boxplot (siehe Abb. 21 auf Seite 117).

6.5.3 Leistungsschwächere Schüler erzielen bessere Ergebnisse, wenn sie sich Zeit nehmen

Bei der erneuten Sichtung des Videomaterials in Bezug auf die Zeit fiel zusätzlich auf, dass sich manche Schüler bei der Bearbeitung des Multiple-Choice-Tests deutlich mehr Zeit nahmen als andere. Der schnellste Schüler war nach 54 Sekunden mit dem Input und dem Multiple-Choice-Test fertig, wohingegen der langsamste Schüler 428 Sekunden benötigte. Das entspricht im Vergleich dem 8-fachen (!) der Zeit. Bei genauere Betrachtung ist zudem aufgefallen, dass sich besonders leistungsschwache Schüler bei der Bearbeitung des Multiple-Choice-Wissenstests deutlich mehr Zeit ließen als leistungsstärkere Schüler. Auch fiel auf, dass leistungsschwächere Schüler – wenn sie den Test so lange bearbeiten konnten, wie sie wollten – tatsächlich viel besser abschritten als erwartet. Hieraus entstand folgende Frage.

Fragestellung: Hat die Bearbeitungszeit bei leistungsschwächeren Schülern einen nachweisbaren Einfluss auf deren Leistung im Multiple-Choice-Test?

Methodisches Vorgehen: Um dieser Frage nachzugehen, wurden alle Videos nachträglich kodiert. Hierfür galt, dass die Bearbeitungszeit beginnt, wenn der Schüler den Input und den Multiple-Choice-Test vom Versuchsleiter erhält. Die Bearbeitungszeit endet, wenn der Schüler den Stift aus der Hand legt und sagt, dass er fertig ist. Zur Beantwortung der Frage, ob die Bearbeitungszeit für das Ergebnis des Tests eine Rolle spielt, wurde im Anschluss eine lineare Regression berechnet. Dabei wurde die Bearbeitungszeit als unabhängige Variable (Prädiktor) und das Testergebnis als abhängige Variable (Kriterium) behandelt.

Ergebnis: Wie in Tab. 31 dargestellt, zeigte das Ergebnis der linearen Regression, dass die Bearbeitungszeit bei guten Schülern keinen signifikanten Vorhersagewert hat. Im Gegensatz dazu stellt die Bearbeitungszeit bei leistungsschwächeren Schülern einen hoch signifikanten Prädiktor für das Ergebnis im Wissenstest dar ($B = .413$, $SE = .005$, $p = .004$, $R^2 = .171$, $df = 45$, mit einer mittleren Effektstärke von $f^2 = .206$ nach Cohen (Cohen 1992)).

Die Rohdaten aller Schüler sind mittels eines Streudiagramms in Abb. 24 dargestellt. Die Aufteilung der Rohdaten in Bezug auf leistungsstärkere und leistungsschwächere Schüler findet sich in Abb. 25 und Abb. 26.

Tab. 31: Ergebnisse der linearen Regression in Bezug auf den Vorhersagewert der Schülerleistung im Multiple-Choice-Test in Abhängigkeit zur Bearbeitungszeit. Signifikante Ergebnisse sind fett gedruckt.

		<i>Nicht standardisierte Koeffizienten</i>		<i>Standardisierte Koeffizienten</i>	<i>T</i>	<i>p</i>
		<i>Regressionskoeffizient B</i>	<i>Standardfehler</i>	<i>Beta</i>		
Alle Schüler	(Konstante)	21.401	0.852		25.102	.000
	Zeit zum Ausfüllen des MC-Tests	.009	.004	.211	2.206	.030
Leistungsstärkere Schüler	(Konstante)	23.110	1.258		18.375	.000
	Zeit zum Ausfüllen des MC-Tests	<.001	.006	-.003	-.021	.983
Leistungsschwächere Schüler	(Konstante)	20.059	1.139		17.606	.000
	Zeit zum Ausfüllen des MC-Tests	.016	.005	.413	3.046	.004

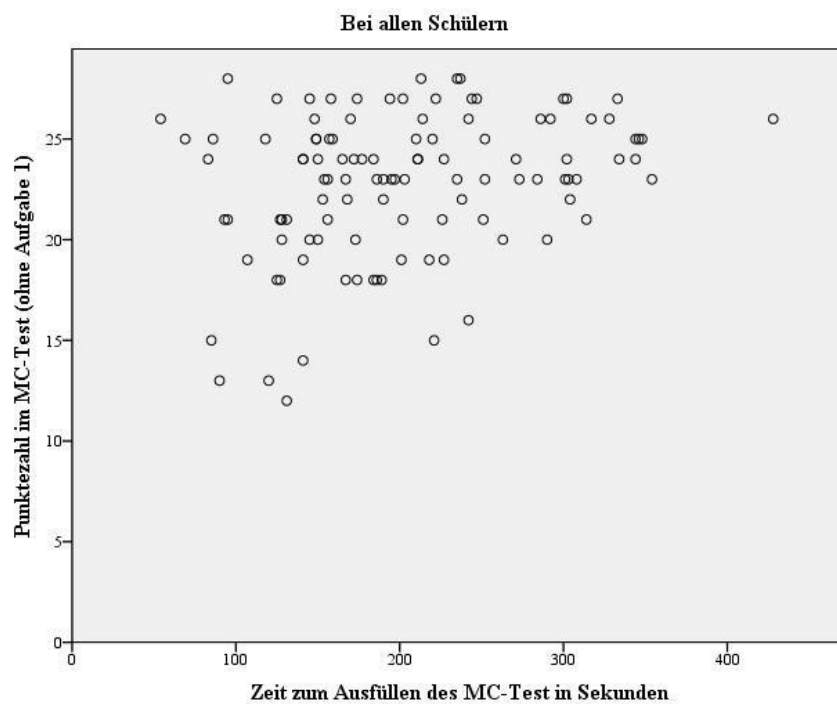


Abb. 24: Streudiagramm der Punktezahl im MC-Test und der Zeit zum Ausfüllen des MC-Tests in Sekunden bei allen Schülern.

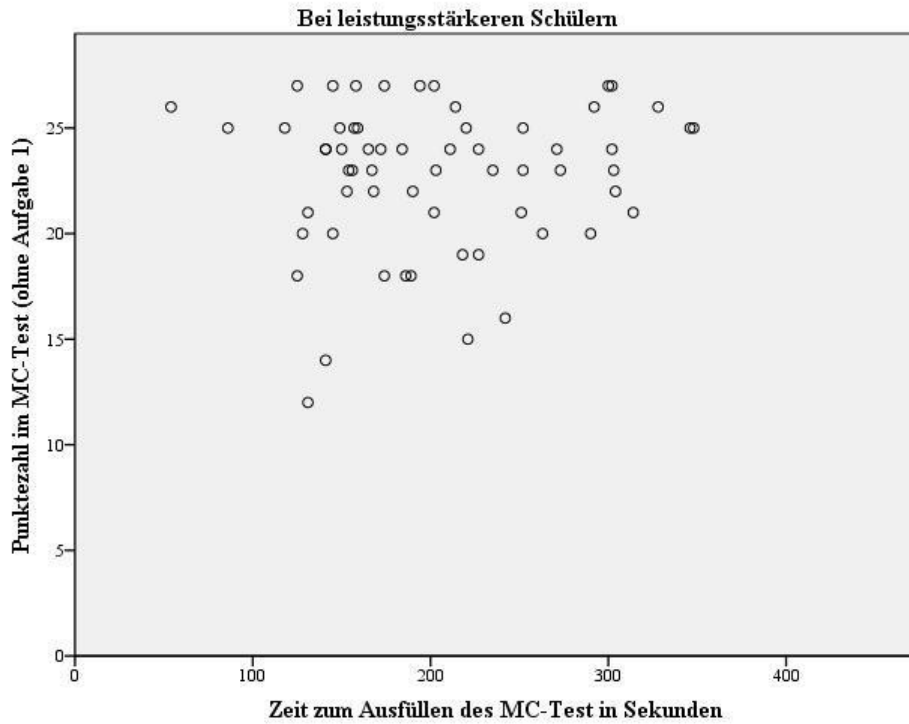


Abb. 25: Streudiagramm der PunktezahI im MC-Test und der Zeit zum Ausfüllen des MC-Tests in Sekunden bei leistungsstärkeren Schülern.

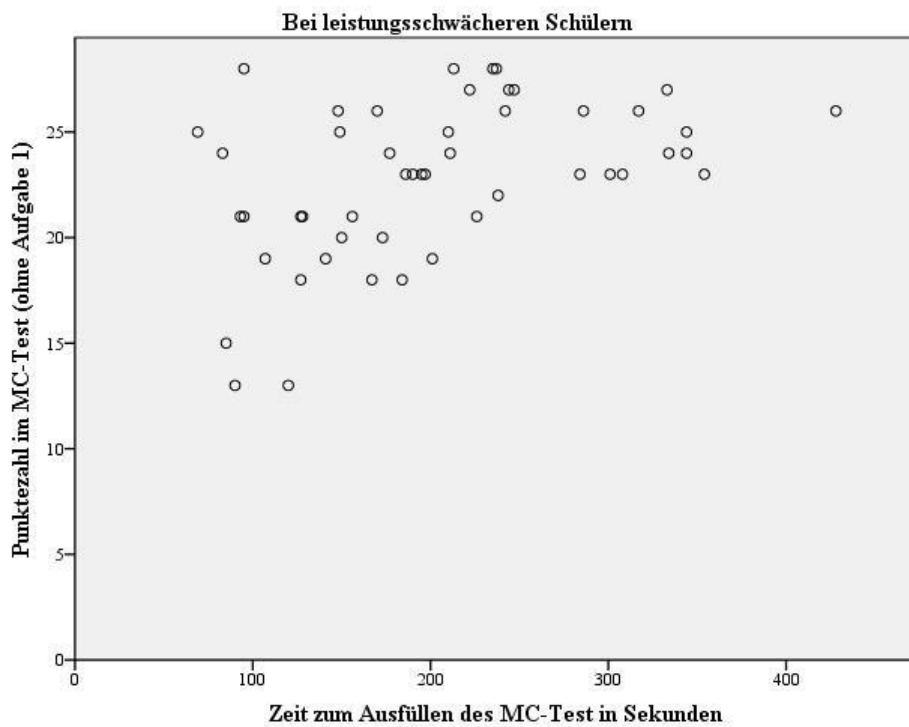


Abb. 26: Streudiagramm der PunktezahI im MC-Test und der Zeit zum Ausfüllen des MC-Tests in Sekunden bei leistungsschwächeren Schülern.

6.5.4 Zeitbedarf bei den Lernaufgaben 1 bis 4

In den Videos wurde die Zeit in Sekunden kodiert, in der die Schüler (n=110) die Lernaufgaben bearbeiteten. Die Zeitmessung begann, als die Schüler die Aufgabenstellung erhielten und endete, wenn die Lösung entweder von den Schülern oder dem Versuchsleiter gesagt worden ist. Über die Daten aller Interviews hinweg wurden die Mittelwerte gebildet und einander gegenübergestellt. Sofern sich hierbei Unterschiede in den Diagrammen zeigten, wurden diese genauer untersucht.

Methodisches Vorgehen: Da es sich bei dem Faktor Zeit um intervallskalierte Messungen handelt, konnten diese mit parametrischen Methoden ausgewertet werden. Für die Suche nach signifikanten Unterschieden zwischen zwei unabhängigen Gruppen in Bezug auf ein Merkmal bietet sich der t-Test an. Zur Berechnung von t-Tests mussten die Daten zuvor aufbereitet werden. Fehlende Werte und Ausreißer in Bezug auf die Zeit wurden per Box-Plot identifiziert und von den Berechnungen ausgeschlossen. Durch die Aufbereitung der Daten reduzierte sich bei der Berechnung die Gesamtanzahl der Schüler von 110 auf 84.

1) **Fragestellung:** Wie viel Zeit wurde für die Lernaufgaben benötigt?

Ergebnis: Im Vergleich wurde für Aufgabe 1 im Durchschnitt am meisten Zeit von den Schülern benötigt und für Aufgabe 2 am wenigsten (siehe Abb. 27). Die größte Spannweite trat bei Aufgabe 4 auf.

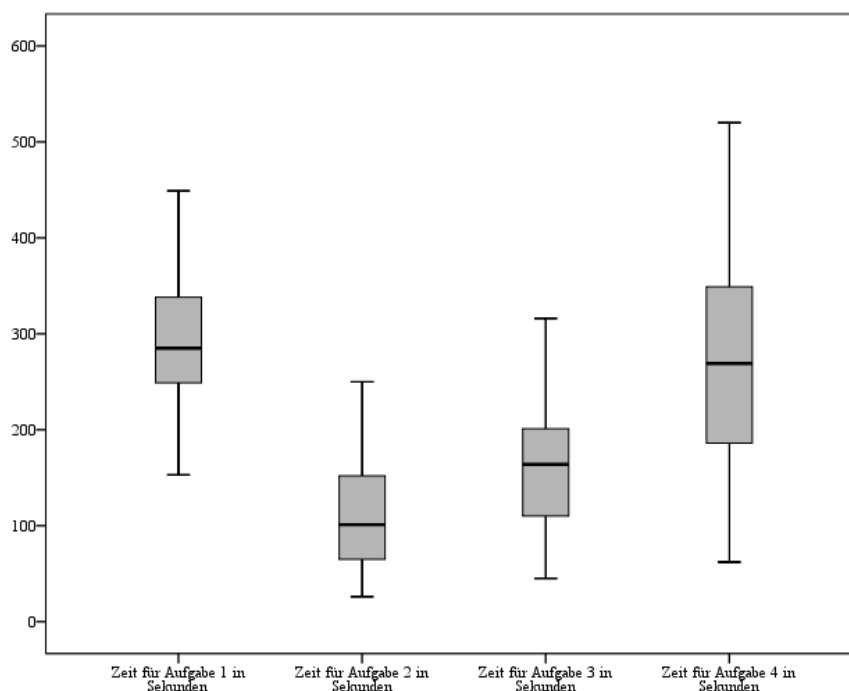


Abb. 27: Bearbeitungszeit aller Schüler (n=84) bei den Lernaufgaben 1 bis 4 in Sekunden.

2) **Fragestellung:** Wie viel Zeit wurde von leistungsstärkeren Schülern im Vergleich zu leistungsschwächeren Schülern für die Aufgaben benötigt?

Ergebnis: Wie in Abb. 28 zu sehen ist, scheinen leistungsschwächere Schüler (n=49) bei allen Aufgaben mehr Zeit für die Lösung zu benötigen als leistungsstärkere Schüler (n=61). Dies konnte für Aufgabe 1 durch einen t-Test bestätigt werden. Hierbei zeigte sich, dass leistungsstärkere Schüler (n=45) marginal signifikant ($p = .076$) weniger Zeit zur Beantwortung von Aufgabe 1 benötigen als leistungsschwächere Schüler (n=36).

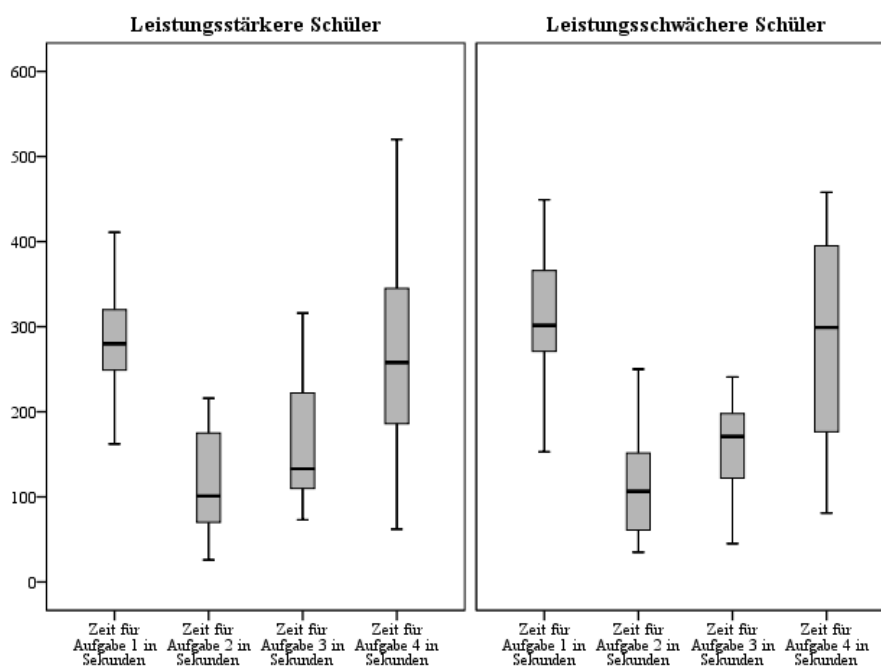


Abb. 28: Bearbeitungszeit für die Aufgaben 1 bis 4 im Vergleich von leistungsstärkeren zu leistungsschwächeren Schülern in Sekunden.

3) **Fragestellung:** Wie viel Zeit wurde von Schülern mit Text (n=55) im Vergleich zu Schülern mit Experten-Concept-Map (n=55) für die Aufgaben benötigt?

Ergebnis: Bis auf Aufgabe 1 scheinen Schüler mit Text bei allen Aufgaben im Durchschnitt mehr Zeit zur Lösung der Aufgabe zu benötigen, als Schüler mit Experten-Concept-Maps (siehe Abb. 29). Diese Vermutung konnte statistisch nicht bestätigt werden, was an den stark überlappenden Streuungsbalken je Aufgabe und Input gut ersichtlich wird.

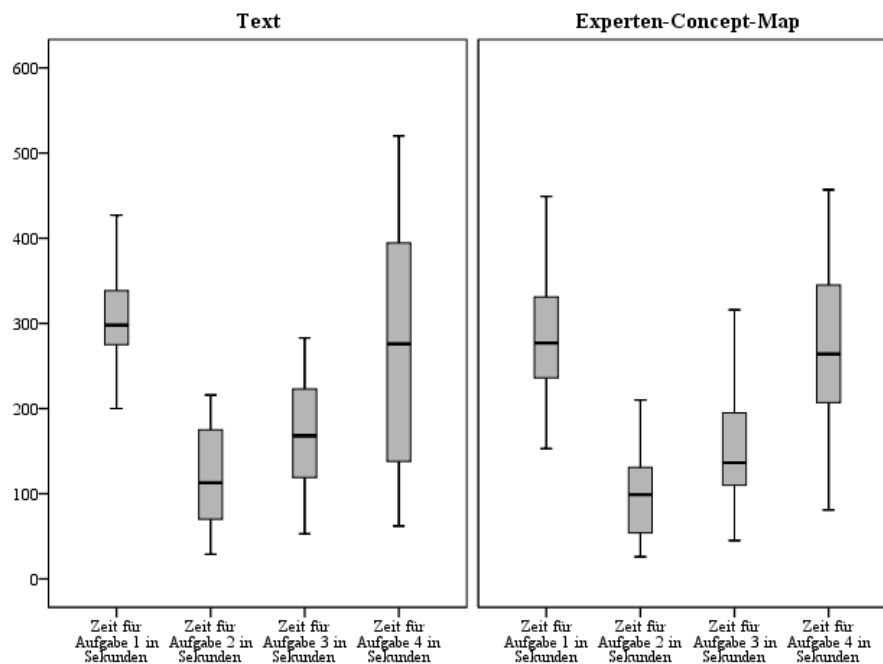


Abb. 29: Bearbeitungszeit für die Aufgaben 1 bis 4 im Vergleich von Schülern mit Text zu Schülern mit Experten-Concept-Map (ECM) in Sekunden.

- 4) **Fragestellung:** Wie viel Zeit wurde von männlichen Schülern im Vergleich zu weiblichen Schülern für die Aufgaben benötigt?

Ergebnis: Bei fast allen Aufgaben scheinen Mädchen (n=40) im Durchschnitt mehr Zeit zur Lösung zu benötigen als Jungen (n=28), außer bei Aufgabe 2 (siehe Abb. 30). Das konnte durch t-Tests bestätigt werden. Mädchen (n=26) und Jungen (n=26) unterscheiden sich in Bezug auf die Bearbeitungszeit hoch signifikant bei den Aufgaben 3 und 4 (jeweils $p = .001$). Besonders eindrücklich ist das an der Spannweite in den beiden folgenden Box-Plots bei Aufgabe 4 zu sehen.

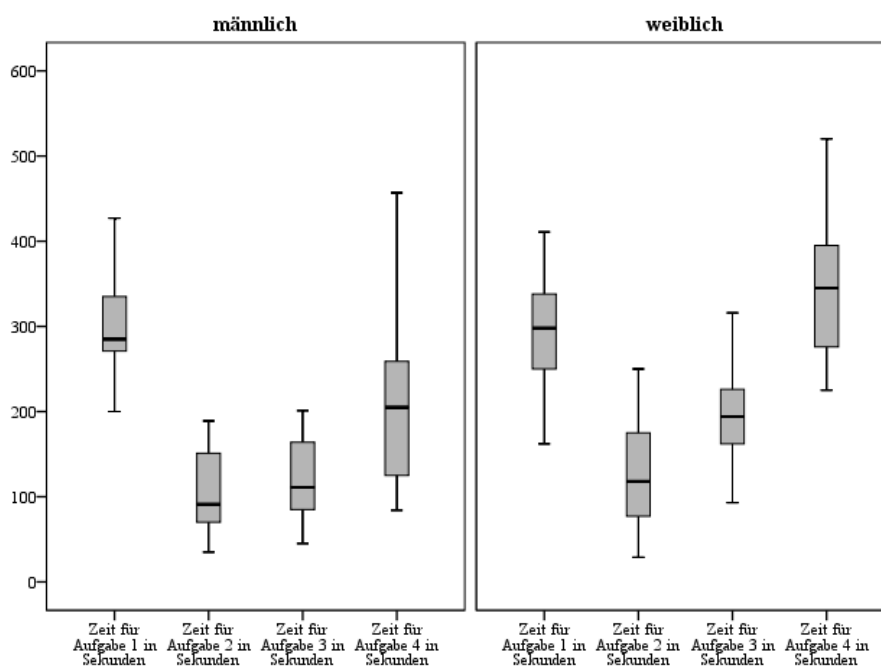


Abb. 30: Vergleich der Bearbeitungszeit von männlichen und weiblichen Schülern bei den Aufgaben 1 bis 4 in Sekunden.

5) **Fragestellung:** Wie viel Zeit wurde von Schülern mit einem guten Ergebnis (n=39) im Multiple-Choice-Test im Vergleich zu Schülern mit einem schlechten Ergebnis im Multiple-Choice-Test (n=42) für die Aufgaben 1 bis 4 benötigt?

Ergebnis: Schüler mit einem guten Ergebnis scheinen über alle Aufgaben hinweg weniger Zeit für die Lösung zu benötigen, als Schüler mit einem schlechten Ergebnis im Multiple-Choice-Test (siehe Abb. 31). Dieser Unterschied erwies sich als nicht signifikant, sowohl für die einzelnen Aufgaben als auch über alle Aufgaben hinweg.

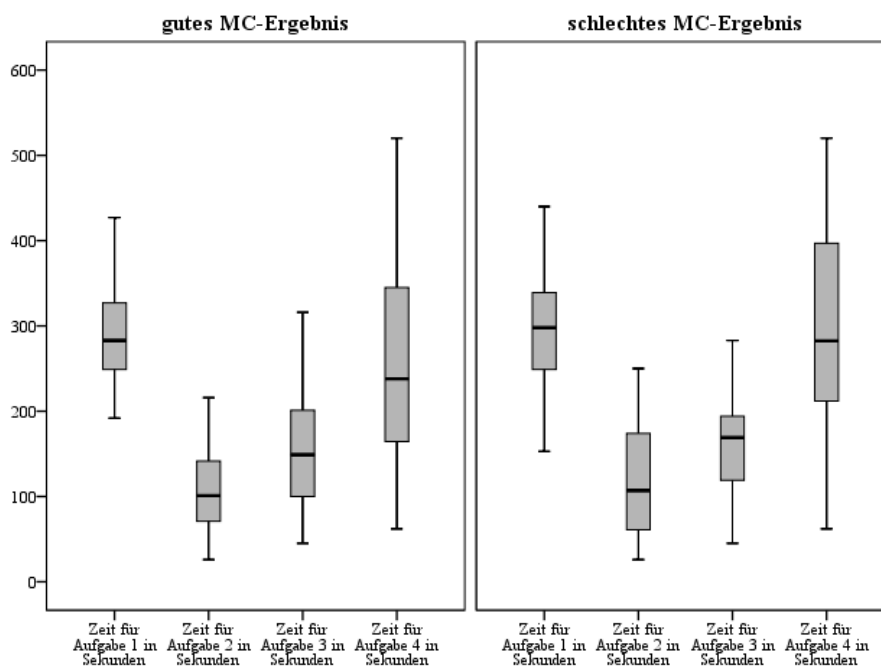


Abb. 31: Bearbeitungszeit für die Aufgaben 1 bis 4 im Vergleich von Schülern mit einem guten Ergebnis zu Schülern mit einem schlechten Ergebnis im Multiple-Choice-Test.

6.6 Einsatz der Hilfen bei den Lernaufgaben

Die vorformulierten Hilfen bilden die notwendigen kognitiven Operationen zum Lösen der Aufgaben ab (siehe Seite 219 ff.). Sie helfen den Schülern, die Aufgabe zu lösen und weisen auf Lernschwierigkeiten der Schüler hin.

6.6.1 Bei Aufgabe 4 wurden die meisten Hilfen benötigt

Fragestellung: Wie viele Hilfen wurden insgesamt bei den Lernaufgaben benötigt?

Methodisches Vorgehen: Zur Dokumentation der eingesetzten Hilfen wurden die Stellen im Video kodiert, an denen der Versuchsleiter in irgendeiner Weise Hilfe oder Unterstützung gegeben hat. In der Regel erfolgte dies durch die Verwendung der farbigen Hilfe-Zettel. In vielen Fällen wurde die Hilfe vom Versuchsleiter auch verbal gegeben oder nur in Teilen, da die Schüler nicht die ganze Hilfe benötigten und der Groschen bereits nach einem Teil der Hilfe gefallen war. Auch diese Hilfen wurden ausgewertet und als ganze Hilfe gezählt.

Ergebnis: Bei der anschließenden Auszählung aller Hilfen kam heraus, dass Schüler bei Aufgabe 2 am wenigsten und bei Aufgabe 4 am meisten Hilfe benötigten (siehe Tab. 32)

Tab. 32: Gesamte Anzahl aller Hilfen pro Aufgabe.

	Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgabe 4
Anzahl der Hilfen	97	34	63	105

Das Ergebnis deckt sich mit der Einschätzung der Schüler, die nach eigener Aussage Aufgabe 2 als leichteste und Aufgabe 4 als schwerste empfanden (siehe Tab. 36 auf Seite 155).

Fragestellung: Welche Hilfen wurden bei welcher Aufgabe wie häufig verwendet?

Methodisches Vorgehen: Die im Video kodierten Stellen mit Hilfen wurden pro Aufgabe sortiert und ausgezählt.

Ergebnis: Tab. 33 zeigt, welche Hilfen pro Aufgabe wie häufig vom Versuchsleiter gegeben wurden. Eine nähere Betrachtung der eingesetzten Hilfen erfolgt unter Tab. 35 auf Seite 147.

Tab. 33: Häufigkeit der gegebenen Hilfen pro Aufgabe. Die am häufigsten eingesetzten Hilfen pro Aufgabe sind fett gedruckt.

	Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgabe 4
Hilfe 1	10	0	0	23
Hilfe 2	2	3	2	11
Hilfe 3	29	5	30	46
Hilfe 4	40	25	19	15
Hilfe 5	16	1	12	10

6.6.2 Qualität der benötigten Hilfe

Gleichermaßen interessant wie die Frage nach der Anzahl der Hilfen ist die Frage nach deren Qualität.

Fragestellung: Welche qualitativen Merkmale besitzen die jeweiligen Hilfen?

Methodisches Vorgehen: Genauso wie es zwischen den Lernaufgaben qualitative Unterschiede gab, gab es diese auch zwischen den Hilfen. Die Hilfen lassen sich in zwei Kategorien unterteilen. Zum einen sind das fachliche Hilfen (FH) und zum anderen sind das Hilfen, die den Schülern dabei helfen, den Arbeitsauftrag oder das Problem besser zu verstehen (APV).

Entsprechend dieser beiden Merkmale wurden alle Hilfen überprüft und den beiden Kategorien zugeordnet. Bei zwei Hilfen war die Zuordnung nicht eindeutig möglich, aus diesem Grund wurden diese in zwei Durchläufen kategorisiert. Bei dem ersten Durchgang durften die Hilfen nur einer Kategorie zugeordnet werden (HART). Beim zweiten Durchgang waren Doppelnennungen möglich (SOFT).

Ergebnis: Bei der harten Trennung der Hilfen kam heraus, dass 26% aller Hilfen gegeben wurden, damit die Schüler die Aufgabe verstehen. Bei der soften Variante stieg dieser Wert auf 37%. Die Kategorisierung in fachliche Hilfen und Hilfen zum Verstehen der Aufgaben sind in Tab. 34 dargestellt, die einer Erweiterung von Tab. 33 entspricht.

Tab. 34: Anzahl und Aufteilung der Hilfen in zwei Kategorien: Aufgaben-Problem-Verständnis (APV) und rein fachliche Hilfe (FH). Hilfen, die sich auf APV und FH beziehen, sind grau hinterlegt.

	Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgabe 4
Hilfe 1	10 APV	0 APV	0 APV	23 APV
Hilfe 2	2 APV	3 APV	2 APV	11 FH/APV
Hilfe 3	29 FH	5 FH	30 FH	46 FH
Hilfe 4	40 FH/APV	25 FH	19 FH	15 FH
Hilfe 5	16 APV	1 APV	12 APV	10 APV

6.7 Schwierigkeiten der Schüler beim Lösen der Aufgaben

Die Trennung in fachliche und nicht-fachliche Hilfen hat gezeigt, dass ein Viertel aller Schüler nicht-fachliche Hilfen zum Lösen der Aufgaben benötigen. Um zu klären, worin genau die fachlichen beziehungsweise nicht-fachlichen Probleme liegen, dient die folgende Fragestellung:

Fragestellung: Welche fachlichen beziehungsweise nicht-fachlichen Hilfen wurden bei den Aufgaben am häufigsten benötigt?

Methodisches Vorgehen: Um genauere Aussagen über die fachlichen und nicht-fachlichen Probleme der Schüler treffen zu können, wurden beide Kategorien in Unterkategorien aufgeteilt. Hierfür wurden alle Hilfen erneut überprüft und entsprechende Subkategorien formuliert. Für den Bereich der fachlichen Hilfe konnten folgende Subkategorien gebildet werden: Funktion der Stäbchen und Zapfen (Hell-Dunkel- und Farbwahrnehmung), Eigenschaften der Stäbchen und Zapfen (Lichtempfindlichkeit), Funktion der Schaltzellen, Aufbau der Netzhaut und Lichteinfall ins Auge. Im Bereich der nicht-fachlichen Hilfe waren dies: Hilfe zum Verstehen der Aufgabenstellung, Hilfe zum genauen Betrachten der Abbildungen, Unterstützung bei der Formulierung der Antwort und einen Denkanstoß geben. Mehr Hilfen in Bezug auf fachliche oder nicht-fachliche Unterstützung konnten nicht identifiziert werden und waren für die Lösung der Aufgaben auch nicht notwendig.

Da sich die Aufgaben ausschließlich um die fünf fachlichen Inhalte gedreht haben, kamen die Hilfen nicht nur in einer Aufgabe vor sondern in mehreren. Um eine Aussage über deren Häufigkeit über alle vier Aufgaben hinweg zu treffen, mussten die jeweiligen Unterkategorien zuerst identifiziert und anschließend zusammengezählt werden. Zum besseren Verständnis dieses Schrittes, sind die Subkategorien in Tab. 35 aufgeführt.

Ergebnis: Die Addition der Unterkategorien von fachlichen und nicht-fachlichen Hilfen über alle vier Aufgaben hinweg ermöglichte ein Ranking der Hilfen innerhalb jeder Kategorie. Dieses Ranking ist absteigend aufgelistet. Die Zahl in Klammern entspricht der Anzahl der Verwendungen der jeweiligen Hilfen im Interview über alle Lernaufgaben hinweg. Folgend kommt das Ranking der fachlichen Hilfen:

- 1) Hilfe zum Aufbau der Netzhaut (91)
- 2) Hilfe zur Lichtempfindlichkeit der Stäbchen und Zapfen (85)
- 3) Hilfe zum Lichteinfall ins Auge (76)
- 4) Hilfe zur Funktion der Zapfen und Stäbchen (53)
- 5) Hilfe zur Funktion der Schaltzellen (40)

Im Folgenden findet sich das Ranking der nicht-fachlichen Hilfen:

- 1) Hilfe zum genauen Betrachten der Abbildungen (42)
- 2) Unterstützung bei der Formulierung der Antwort (39)
- 3) Hilfe zum Verstehen der Aufgabenstellung oder des Problems (33)
- 4) Das Geben von Denkanstößen (16), Beispiele hierzu zeigt Tab. 35.

Alle in der folgenden Tabelle genannten Hilfen werden in Kapitel 9.2.7 ab Seite 242 aufgeführt.

Tab. 35: In dieser Tabelle wird in jeder Zelle für jede Hilfe zuerst deren Häufigkeit in den Interviews aufgeführt. Die anschließende Abkürzung gibt Auskunft über die Hauptkategorie der Hilfe (APV = Aufgaben-Problem-Verständnis; FH = fachliche Hilfe). Der darauffolgende Satz beschreibt die jeweilige Subkategorie der Hilfe. Die letzte Zahl gibt an, wie viel Prozent der Schüler diese Hilfe im Interview beansprucht haben.

	Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgabe 4
Hilfe 1	10 APV-Hilfe zum Verstehen der Aufgabe oder des Problems (18%)	0 APV-Hilfe zum Verstehen der Aufgabe oder des Problems (0%)	0 APV-Hilfe zum Verstehen der Aufgabe oder des Problems (0%)	23 APV-Hilfe zum Verstehen der Aufgabe oder des Problems (42%)
Hilfe 2	2 APV-Hilfe zum Betrachten der Abbildungen (4%)	3 APV-Denkanstoß <i>(Was ist der Unterschied zwischen Tag und Nacht?)</i> (5%)	2 APV-Denkanstoß <i>(Was könnte helfen, damit man nachts besser sieht?)</i> (4%)	11 FH/APV-Eigenschaften der Stäbchen und Zapfen (20%)
Hilfe 3	29 FH-Funktion der Stäbchen und Zapfen (53%)	5 FH-Funktion der Stäbchen und Zapfen (9%)	30 FH-Lichteinfall / Aufbau Netzhaut / Eigenschaften Stäbchen und Zapfen (55%)	46 FH-Lichteinfall / Aufbau der Netzhaut (84%)
Hilfe 4	40 FH/APV-Funktion der Schaltzellen / Hilfe zum genaueren Betrachten der Abbildungen (72%)	25 FH-Eigenschaften der Stäbchen und Zapfen (45%)	19 FH-Eigenschaften der Stäbchen und Zapfen / Funktion der Stäbchen und Zapfen (35%)	15 FH-Aufbau der Netzhaut (27%)
Hilfe 5	16 APV-Unterstützung bei der Formulierung der Antwort (29%)	1 APV-Unterstützung bei der Formulierung der Antwort (2%)	12 APV-Unterstützung bei der Formulierung der Antwort (22%)	10 APV-Unterstützung bei der Formulierung der Antwort (18%)

Das Ranking der Hilfen lässt sich natürlich nicht nur über alle Aufgaben hinweg durchführen, sondern auch für jede Aufgabe einzeln. Durch das Ranking der Hilfen innerhalb einer Aufgabe lassen sich für Lehrkräfte direkte Handlungsanweisungen für den konkreten Unterricht ableiten.

Setzen Lehrkräfte zum Beispiel Aufgabe 1 in ihrem Unterricht ein, so werden mit hoher Wahrscheinlichkeit die meisten Schüler Probleme beim Verstehen der Abbildungen und der Funktion der Schaltzellen haben. Wenn Lehrkräfte das wissen, können sie im Vorfeld darauf eingehen und dieses Problem durch zusätzliche Erklärungen, Anweisungen oder Material lösen. Für Aufgabe 2 wäre dies nicht nur die Betonung der Lichtwahrnehmung der Sehsinneszellen, sondern auch auf der unterschiedlichen Lichtempfindlichkeit. Bei Aufgabe 3 und 4 wäre besonders der Aufbau der Netzhaut in Verbindung mit dem Lichteinfall ins Auge im Unterricht zu behandeln.

Häufigkeit der genutzten Hilfen bei Schülergruppen

Neben den bereits aufgeführten Aussagen zu den Hilfen und den damit verbundenen Verständnisschwierigkeiten aller Dyaden in dieser Studie, lassen sich anhand der Daten noch präzisere Aussagen über die Probleme der Schülergruppen treffen. Genauso wie bei dem Vergleich der Antwortqualität zwischen unterschiedlichen Schülergruppen (siehe 6.4 ab Seite 129), ist aufgrund der teilweise stark variierenden Größe der nachfolgenden Gruppen und durch die Nominalskalierung der Daten ein Vergleich der Mittelwerte nicht zweckdienlich.

Dementsprechend wurde für die nachfolgende Darstellung der Daten der Prozentsatz für die Häufigkeit des Einsatzes jeder Hilfe gebildet und in den Abbildungen graphisch dargestellt. Entsprechend der Skalierung der Daten wurden bei dem Vergleich der Schülergruppen alle Werte mittels Chi-Quadrat-Test nach Pearson berechnet. Hieraus resultierende und statistisch signifikante Werte sind in den Abbildungen mit einem schwarzen Rahmen gekennzeichnet und werden darüber hinaus im Fließtext genannt. Bevor die Ergebnisse der Gruppenvergleiche dargestellt werden, wird als Referenzgröße zuerst die prozentuale Häufigkeit der Hilfen für alle Dyaden aufgeführt.

Aus Gründen der Lesbarkeit wurden in den folgenden Balkendiagrammen Kürzel für die Hilfen verwendet. Die Abkürzung *AIH1* steht für *Aufgabe 1 Hilfe 1*, das Kürzel *AIH2* steht für *Aufgabe 1 Hilfe 2* und so weiter. Die ausformulierten Hilfen werden in Kapitel 9.2.7 ab Seite 240 aufgeführt.

1) **Fragestellung:** Wie häufig wurden welche Hilfen von den Dyaden im Interview benötigt?

Ergebnis: Bei allen Dyaden (n = 57) wurde die Hilfe Nummer 3 bei Aufgabe 4 prozentual am häufigsten benötigt. Gefolgt von Hilfe 3 und Hilfe 4 bei Aufgabe 1. Bei der Betrachtung von Abb. 32 wird deutlich, dass die Hilfen bei den Aufgaben ganz unterschiedlich zum Einsatz gekommen sind.

Abb. 32 ist im Prinzip nichts anderes als eine Übersetzung der bereits in Tab. 33 aufgeführten Anzahl der jeweiligen Hilfen pro Aufgabe. Die prozentuale Darstellung im Balkendiagramm ermöglicht jedoch nicht nur einen direkten Vergleich innerhalb der jeweiligen Schülergruppen, sondern auch zwischen allen Gruppen dieser Studie.

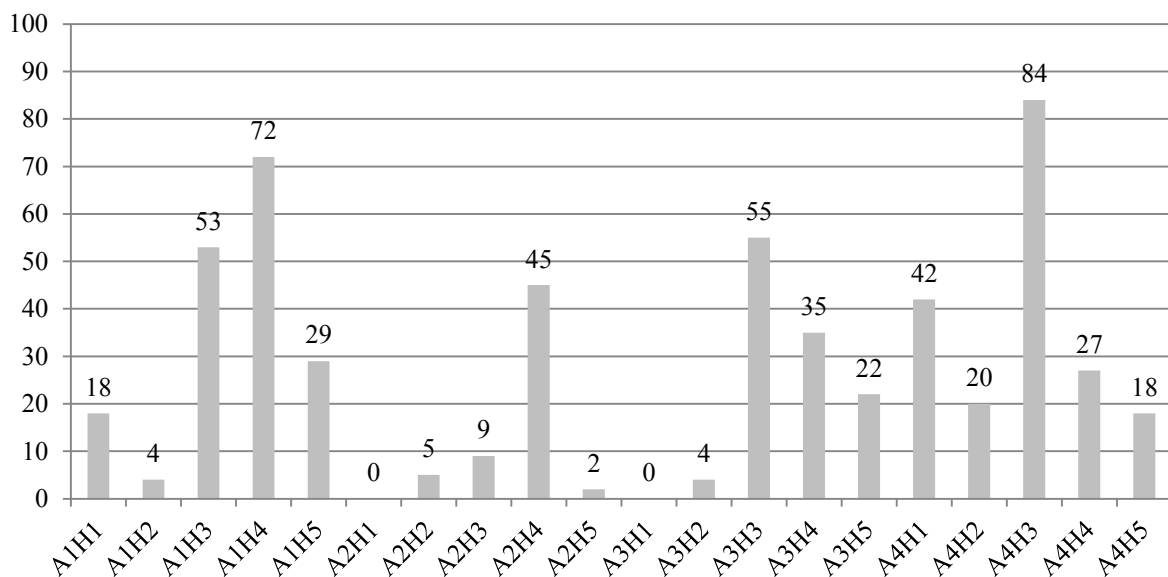


Abb. 32: Prozentuale Häufigkeit des Einsatzes der Hilfen bei allen Dyaden. Bei der Beschriftung der x-Achse steht A für Aufgabe und H für Hilfe. Der Balken über A1H1 stellt demnach die prozentuale Häufigkeit von Hilfe 1 bei Aufgabe 1 bei allen Interviews dar (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).

2) **Fragestellung:** Welche Hilfen benötigen leistungsschwächere Dyaden im Vergleich zu leistungsstärkeren Dyaden?

Ergebnis: Wie zu erwarten, benötigten leistungsschwächere Dyaden (n=24) in der Regel mehr Hilfe als leistungsstärkere Dyaden (n=33). Dies galt besonders für Aufgabe 1 und 2 (siehe Abb. 33). Bei der anschließenden Überprüfung der Signifikanz der Unterschiede, lieferte der Chi-Quadrat-Test nach Pearson einen marginalen Unterschied zwischen leistungsstärkeren und leistungsschwächeren Schülern bei Aufgabe 1 Hilfe 4 ($p = .064$), einen hoch signifikanten Unterschied bei Aufgabe 1 Hilfe 5 ($p = .002$) und einen signifikanten Unterschied bei Aufgabe 2 Hilfe 2 ($p = .037$). Dies bedeutet, dass leistungsschwächere Schüler im Vergleich zu leistungsstärkeren Schülern besonders bei der Funktion der Schaltzellen und dem genauen Betrachten von Abbildungen mehr Probleme haben. Dieses Ergebnis deckt sich mit den Schwierigkeiten der gleichen Schülergruppe beim Beschreiben der Abbildungen nach der Auseinandersetzung mit dem differenzierenden Input (siehe 6.8.6.2 auf Seite 168). Ferner machten ihnen die Formulierung der Antwort bei Aufgabe 1 sowie das Verstehen des Problems bei Aufgabe 2 mehr Schwierigkeiten als leistungsstärkeren Schülern. Bei dem Vergleich der prozentualen Häufigkeit zwischen leistungsstärkeren und leistungsschwächeren Schülern wird auch deutlich, dass signifikante Unterschiede nur bei Aufgabe 1 und 2 auftreten. Bei Aufgabe 3 und 4 liegen leistungsstärkere und leistungsschwächere Dyaden in Bezug auf benötigte Hilfe nahe beieinander. So benötigen sowohl leistungsschwächere als auch leistungsstärkere Schüler bei A4H3 gleichermaßen häufig Hilfe zum Lichteinfall ins Auge und zum Aufbau der Netzhaut.

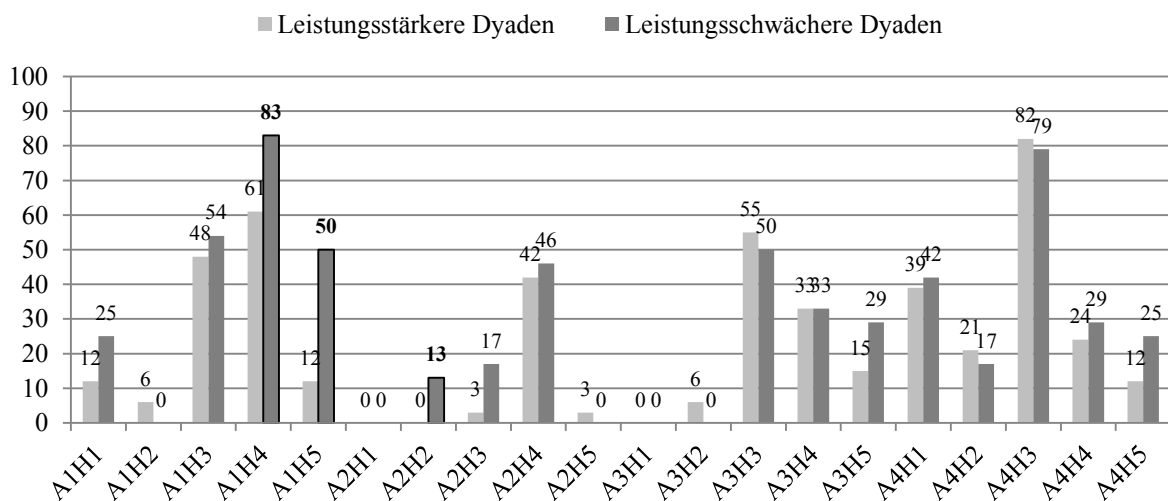


Abb. 33: Vergleich der prozentualen Häufigkeit der Hilfen zwischen leistungsstärkeren und leistungsschwächeren Dyaden. Marginale und signifikante Unterschiede sind durch einen schwarzen Rahmen gekennzeichnet (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).

3) **Fragestellung:** Wie viel prozentuale Hilfe benötigen Schülerdyaden mit Text im Vergleich zu Schülerdyaden mit Experten-Concept-Map als Input?

Methodisches Vorgehen: Hierbei wurde genauso vorgegangen wie bei dem vorherigen Punkt mit dem Unterschied, dass die Hilfen nicht nach Leistungsgruppe aufgeteilt wurden, sondern nach Schülerdyaden mit Text (n=28) und Schülerdyaden mit Experten-Concept-Map (n= 29) als Input.

Ergebnis: Die benötigten Hilfen zwischen den beiden Inputgruppen unterscheiden sich kaum. Der Input scheint für die Qualität der Schülerantworten bei den Lernaufgaben keinen sichtbaren Einfluss zu haben (siehe Abb. 34). Außer bei Aufgabe 3 Hilfe 5 konnte kein signifikanter Unterschied zwischen den Gruppen gefunden werden ($p = .044$).

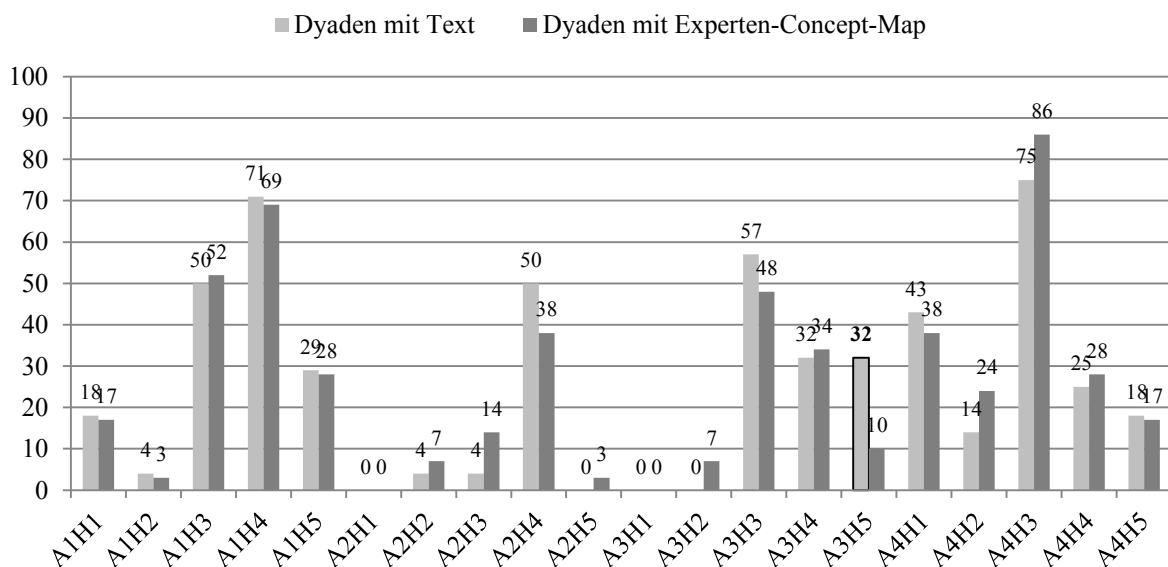


Abb. 34: Vergleich der prozentualen Häufigkeit der Hilfen zwischen Dyaden mit Text und Dyaden mit Experten-Concept-Map. Signifikante Unterschiede sind durch einen schwarzen Rahmen gekennzeichnet (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).

4) **Fragestellung:** Welche Hilfen benötigen Dyaden mit Mädchen im Vergleich zu Dyaden mit Jungen?

Methodisches Vorgehen: Die Trennung nach dem Geschlecht erfolgte im Prinzip genauso wie bei den beiden vorherigen Punkten, allerdings konnten die Gruppen nicht genauso getrennt werden, da es sowohl gleichgeschlechtliche als auch gemischtgeschlechtliche Dyaden gab. Um die Daten auf Geschlechterunterschiede zu untersuchen, konnten nur die gleichgeschlechtlichen Dyaden für die Auswertung herangezogen werden. Das sind bei Jungen 14 (n) Dyaden und bei den Mädchen 20 (n).

Ergebnis: Bei den Aufgaben 1 und 2 scheinen die Jungen etwas mehr Hilfe als die Mädchen zu benötigen. Dies ändert sich bei Aufgabe 3 und ganz besonders bei Aufgabe 4. Hier scheinen die Mädchen deutlich mehr Hilfe zu benötigen als die Jungen (siehe Abb. 35). Bei der nachträglichen Überprüfung der Unterschiede mittels Chi-Quadrat-Test, konnten statistisch signifikante Unterschiede zwischen den Gruppen bei den folgenden Hilfen ermittelt werden: Bei Aufgabe 3 Hilfe 4 ($p = .017$) und bei Aufgabe 4 Hilfe 1 ($p = .050$), Hilfe 3 ($p = .007$) und Hilfe 4 ($p = .033$). Konkret bedeutet dies, dass die gleichgeschlechtlichen Mädchendyaden in dieser Studie mehr Schwierigkeiten bei Aufgabe 3 mit den Eigenschaften und der Funktion der Stäbchen und Zapfen haben. Ferner haben sie bei Aufgabe 4 auch mehr Schwierigkeiten mit der Aufgaben- und Problemstellung, dem Lichteinfall ins Auge und dem Aufbau der Netzhaut.

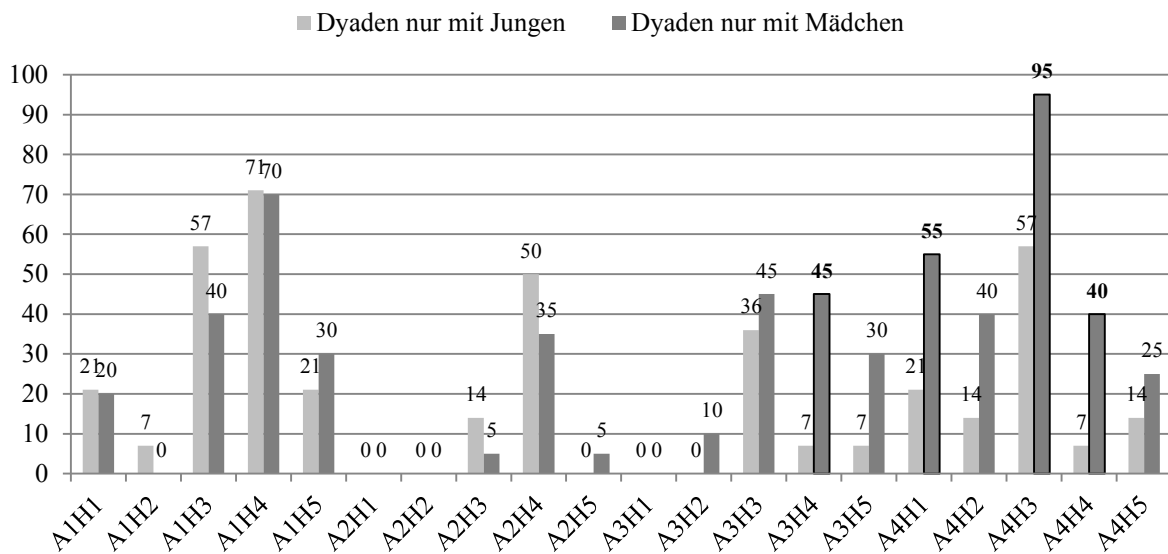


Abb. 35: Vergleich der prozentualen Häufigkeit der Hilfen zwischen Dyaden mit Mädchen und Dyaden mit Jungen. Bei Aufgabe 3 und besonders bei Aufgabe 4 geht die Häufigkeit der beanspruchten Hilfen zwischen den Geschlechtern stark auseinander (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).

5) **Fragestellung:** Welche Hilfen benötigen Dyaden mit einem guten Ergebnis im Multiple-Choice-Test im Vergleich zu Dyaden mit einem schlechten Ergebnis?

Methodisches Vorgehen: Hierfür wurden alle Schülerdyaden entsprechend ihres Ergebnisses im Multiple-Choice-Test aufgeteilt und deren Ergebnisse verglichen.

Ergebnis: Tendenziell scheinen Dyaden mit einem schlechten Ergebnis im Multiple-Choice-Test ($n=27$) mehr Hilfen zu benötigen als Schüler mit einem guten Ergebnis ($n=30$). Dies wird vor allem bei Aufgabe 1 deutlich (siehe Abb. 36). Signifikante Unterschiede zeigten sich in Aufgabe 1 Hilfe 3 ($p = .024$) und Hilfe 5 ($p = .043$), sowie marginale Unterschiede bei Aufgabe 1 Hilfe 4 ($p = .087$) und Aufgabe 2 Hilfe 4 ($p = .091$).

Das Resultat ist dem Vergleich zwischen leistungsstärkeren und leistungsschwächeren Dyaden sehr ähnlich (siehe Seite 150 ff.). Ferner deckt sich dieses Ergebnis mit den Schwierigkeiten von Schüler mit einem schlechten Ergebnis im Multiple-Choice-Test (siehe Seite 168). Auch bei der Gruppe mit einem schlechten Ergebnis im Multiple-Choice-Test scheint es so zu sein, dass sich ein Nichtverstehen beim Input (egal ob Text oder der Experten-Concept-Map) auf das Verstehen der Abbildungen überträgt. Dies zeigt sich gerade bei Aufgabe 1, weil die Lösung dieser Aufgabe explizit die Verbindung des Inputs mit den Abbildungen erfordert.

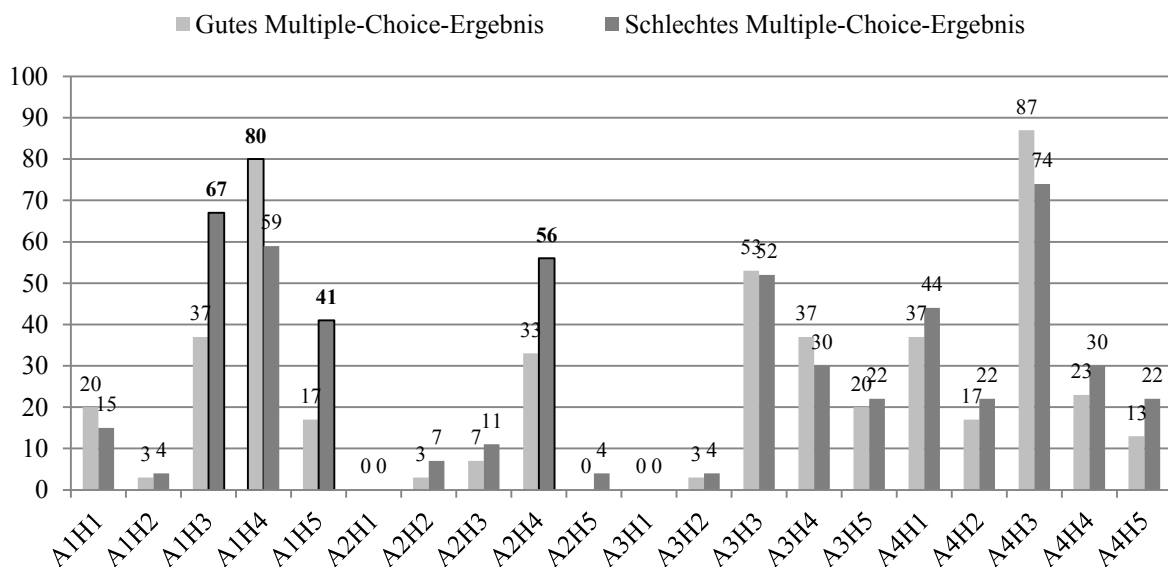


Abb. 36: Vergleich der prozentualen Häufigkeit der Hilfen bei Schülern mit einem guten Ergebnis im Multiple-Choice-Test und bei Schülern mit einem schlechten Ergebnis im Multiple-Choice-Test. Signifikante und marginal signifikante Unterschiede zwischen den Gruppen sind mit einem schwarzen Rahmen markiert (Der Inhalt der jeweiligen Hilfen befindet sich in Tab. 35 auf Seite 147, die ausformulierten Hilfen in Kapitel 9.2.7 ab Seite 242).

6.8 Weitere Fragestellungen

Im Gegensatz zu den vorherigen Unterkapiteln liegt der Fokus der folgenden Abschnitte weniger auf quantitativen, sondern eher auf qualitativen Ergebnissen. Dies beinhaltet unter anderem die Einschätzung des Schwierigkeitsgrades der Aufgaben durch die Schüler, die Qualität der Schülerantworten bei den Aufgaben und bemerkenswerte Erkenntnisse aus der Erprobung der Materialien. Die Gliederung der Unterkapitel folgt dabei den Kapiteln davor.

6.8.1 Einschätzung der Lernaufgaben durch die Schüler

Durch die Erhebung der Qualität der Schülerantworten (siehe Seite 169 ff.) ließen sich weitere Aussagen zu den Aufgaben treffen. Zum Beispiel: Welche der vier Aufgaben wurde von den Schülern am besten gelöst? Rein deskriptiv wurde Lernaufgabe 2 am besten beantwortet und Aufgabe 1 am schlechtesten (siehe Abb. 37 auf Seite 162), doch wie sehen das die Schüler?

Bei dem folgenden Abschnitt handelt es sich um einen Perspektivwechsel da nun die Schüler die selbst bearbeiteten und verwendeten Aufgaben und Materialien rückblickend einschätzen und bewerten konnten.

Fragestellung: Welche der vier Aufgaben empfinden Schüler am leichtesten, am schwersten, am interessantesten?

Methodisches Vorgehen: Hierfür wurden alle Schüler am Ende der Interviews gefragt, welche der Aufgaben sie als am einfachsten, am schwersten und welche sie als am interessantesten empfunden haben. Im Interview wurde akzeptiert, wenn Schüler hierzu keine eigene Meinung hatten oder diese nicht sagen wollten. Darüber hinaus waren im Zweifelsfall auch Mehrfachnennungen möglich. Alle Nennungen wurden addiert und pro Aufgabe (1-4) und Kategorie (leicht, schwer, interessant) summiert.

Ergebnis: Die geäußerten Schülermeinungen sind eindeutig ausgefallen (siehe Tab. 36). Mit großem Abstand empfanden die Schüler Aufgabe 2 am leichtesten und Aufgabe 4 am schwersten. Bemerkenswerterweise stufte der Großteil der Schüler die schwerste Aufgabe aber gleichzeitig auch als die interessanteste Aufgabe ein.

Neben der eindeutigen Einschätzung aller Schüler zusammen, trat bei dem Vergleich der Geschlechter ein weiterer interessanter Punkt auf: Mädchen und Jungen empfanden Aufgaben unterschiedlich schwer. So schätzten Mädchen Aufgabe 1 als leicht und Aufgabe 3 als schwierig ein, wohingegen Jungen Aufgabe 1 als schwierig empfinden und Aufgabe 3 als leicht. Die Einschätzung der beiden Geschlechter in Bezug auf den

Schwierigkeitsgrad dreht sich bei diesen beiden Aufgaben um. Bei dem Interessanzgrad der Aufgabe gibt es hingegen keinen nennenswerten Geschlechterunterschied (siehe Tab. 37).

Tab. 36: Häufigkeit der Einschätzungen aller Schüler (n=110) zum Schwierigkeits- und Interessanzgrad der Aufgaben. Die Schüler nannten jeweils die einfachste, schwerste und interessanteste Aufgabe. Doppelnennungen und Enthaltungen waren erlaubt.

	Aufg. 1	Aufg. 2	Aufg. 3	Aufg. 4
Einfachste Aufgabe	32	58	17	14
Schwerste Aufgabe	16	11	22	72
Interessanteste Aufgabe	6	24	21	64

Tab. 37: Häufigkeit der Schülereinschätzungen zum Schwierigkeits- und Interessanzgrad der Aufgaben in Bezug auf das Geschlecht (♂ n=49 und ♀ n=61). Starke Unterschiede zwischen den Geschlechtern sind fett gedruckt.

	Geschlecht	Aufg. 1	Aufg. 2	Aufg. 3	Aufg. 4
Einfachste Aufgabe	♂	10	24	12	7
	♀	22	34	5	7
Schwerste Aufgabe	♂	11	4	8	29
	♀	5	7	14	43
Interessanteste Aufgabe	♂	2	11	10	28
	♀	4	13	11	36

Verbale Einschätzung der Aufgaben 1 bis 4 durch Schüler

Fragestellung: Wie begründen Schüler, was eine Lernaufgaben aus ihrer Sicht leicht, schwer oder interessant gemacht hat?

Methodisches Vorgehen: Nachdem die Schüler ihre Meinung zu den Aufgaben geäußert hatten, wurden sie aufgefordert, ihre Aussage auch zu begründen: *Warum empfindest du die Aufgabe als am schwersten, am einfachsten oder am interessantesten?* Bei der Analyse der Videos ist schnell aufgefallen, dass zwar alle Schüler eine Meinung zu den Aufgaben hatten, allerdings sehr viele bei der Begründung ihrer Aussagen Schwierigkeiten hatten. Sehr häufig ging die Begründung nicht über ein „ich weiß“ nicht“ oder „es ist halt so“ hinaus. Ein Schüler sagte zum Beispiel: „ach, keine Ahnung... für uns war die halt schwer!“. Oft endete die Begründung auch in der Wiederholung der Aufgabenstellung oder der Lösung der Aufgabe.

Die qualitative Auswertung der Schülerantworten erfolgte immer auf die gleiche Weise:

- 1) Entsprechend der Fragestellung wurden die relevanten Schüleraussagen in den Videos markiert und transkribiert.
- 2) Anschließend wurden Kategorien gebildet. Hierbei wurde datengestützt und induktiv nach dem Grounded Theory Ansatz von Aussage zu Aussage vorgegangen (Glaser und Holton 2004). Schritt für Schritt wurden zuerst grobe Kategorien gebildet, die explizit auf dem Material basieren, also wörtlich in der Aussage zu lesen waren. Zur Erhöhung der Objektivität galt bei dieser Kategorienbildung die Prämisse, dass nur das zählt, was ein Schüler tatsächlich sagt und nicht das, was er meinen könnte.
- 3) Im ersten Durchlauf wurden die Schüleraussagen den daraus resultierten Kategorien zugeordnet.
- 4) Innerhalb jeder Kategorie wurden im nächsten Schritt Subkategorien gebildet, die eine weitere Zuordnungsebene und Hierarchisierung innerhalb jeder Kategorie erlaubte.
- 5) Im letzten Schritt wurden die Kategorien entsprechend ihrer Anzahl an Nennungen geordnet und ausgezählt. Die Auszählung diente der Gewichtung der qualitativen Aussagen und lässt Aussagen über deren Häufigkeit zu.

Die folgenden drei Abschnitte sind dementsprechend jeweils vom häufigsten Merkmal zum seltensten Merkmal sortiert.

6.8.2 Was eine Aufgabe nach Schülermeinung einfach macht

Auf die Frage, was eine Aufgabe einfach macht, antwortete die Mehrheit der Schüler mit den **Lernmaterialien (n=21)**, die die Aufgaben begleiteten. Die Vereinfachung der Aufgaben durch die Zugabe von zusätzlichen Lernmaterialien begründeten die Schüler damit, dass die Antworten meistens schon im Material stecken und man sie nur finden müsse. Das Finden der Antworten erfordere nach Schülermeinung gutes und aufmerksames Durcharbeiten der Materialien. Als explizite Unterstützung nannten die Schüler den **Text (n=15)** und die **Abbildungen (n=6)**. Den Text vor allem daher, weil er die nötigsten Informationen enthält und auch schon ein Teil der Erklärung liefert. Die Abbildungen punktetten dadurch, dass man diese nur betrachten und vergleichen muss und somit die Lösung direkt vor Augen hat. Manche Schüler sagten sogar, dass sie ohne die Abbildungen überhaupt nichts hätten machen können. Eine Person drückte aus, dass sie ohne das Augenmodell nicht verstanden hätte, wo der gelbe Fleck liegt. Interessanterweise hat keiner der Schüler die Experten-Concept-Map als erleichterndes Merkmal erwähnt.

Das zweithäufigste Merkmal für einfache Aufgaben nach Schülermeinung ist die **Reihenfolge der Aufgaben (n=5)**. Hierbei traten unterschiedliche Meinungen auf, die verschiedene Aspekte der Aufgabenreihenfolge ansprechen. Zwei der Schüler fanden zum Beispiel, dass die erste Aufgabe am schwersten war, weil ihnen der Ablauf mit den Aufgaben mit gestuften Hilfen bisher unbekannt war und sie sich erst daran gewöhnen mussten. Das hat weniger mit den Aufgaben oder deren Reihenfolge an sich zu tun, sondern mit der Vertrautheit der Methode. Ein Anderer hingegen sagte, dass die erste Aufgabe eher schwierig gewesen sei, weil man inhaltlich und fachlich noch überhaupt nichts wusste und man sich das Wissen erst im Laufe der Aufgaben angeeignet hat („da haben wir das Nötigste für die anderen Aufgaben gelernt“). Das bedeutet aber nicht, dass die letzte Aufgabe die einfachste sei, denn: „wenn man die anderen Aufgaben davor hatte, ist die vierte Aufgabe einfach, sonst nicht!“. Zwei andere Schüler wiederum fanden, dass die Aufgaben immer schwieriger wurden. Hierfür gaben sie zwar keine Begründung an, allerdings decken sich ihre Aussagen mit der intendierten Staffelung der Aufgaben bezüglich ihres Schwierigkeitsgrades bei der Planung der Studie (siehe Seite 112 ff.).

Als weiteren Punkt für die Erleichterung von Aufgaben nannten Schüler den **Bekanntheitsgrad des Kontextes (n=3)**, in dem sich eine Aufgabe bewegt. Zusätzlich wurde **Vorerfahrung (n=3)** zu einem biologischen Phänomen als erleichternd bezeichnet. Die Schüler sagten, dass es für sie viel leichter sei, eine problemorientierte Aufgabe zu lösen, wenn sie das Problem selbst kennen, es bereits persönlich erlebt oder gesehen haben, oder sie sich dies gut vorstellen können. Auf den letzten Rängen der Häufigkeit nennen Schüler die **Aufgabenstellung (n=2)**. Hier legen sie Wert auf so wenig Aufgabentext wie möglich, am besten „kurz und knackig“.

6.8.3 Was eine Aufgabe nach Schülermeinung schwierig macht

Die absolute Mehrheit der Schüler sieht die Schwierigkeit von Aufgaben in deren **Anforderungen an die Schüler (n=16)**, die ganz besonders durch das **Nachdenken (n=14)** entstehen. Auch wenn sie es nicht auf diese Art ausdrückten, so wird es durch deren Gebrauch von Adjektiven und Umschreibungen klar. Zur Verdeutlichung sind einige davon im Folgenden aufgeführt: „man kommt nicht direkt auf die Lösung...es ist knifflig...man muss mitdenken...man muss nachdenken...man muss umdenken...man muss um die Ecke denken...man muss es sich selbst zusammensetzen“. Da sich diese Beschreibung sehr stark mit dem deckt, was eine Aufgabe nach Schülermeinung auch interessant macht, wird dieser Punkt ab Seite 158 vertieft und im Folgenden vorerst auf Merkmale eingegangen, die sich nach Schülermeinung direkt auf die Schwierigkeit von Aufgaben beziehen.

Zwei weitere erschwerende Merkmale sind die Masse an **Aufgabentext (n=6)** und die Masse an **Aufgabeninhalt (n=3)**. Schon bei der Durchführung der Interviews ist aufgefallen, dass viele Schüler eine Abneigung gegenüber der Länge des Aufgabentextes hatten. Dies wurde besonders beim Auffalten der Aufgabe 4 deutlich. Alleine der Anblick des Textes führte zu ablehnenden Reaktionen. Viele Schüler zeigten ihren Unmut über den Text sowohl verbal als auch nonverbal. Auf die Frage, warum das so sei, antworteten Schüler, dass viel Text schwierig für sie sei oder dass sie sich mit zunehmender Textlänge weniger konzentrieren könnten. Eine Schülerin sagte sogar: „ich hatte einfach keine Lust auf Lesen heute“. Auf die Frage, was genau einen längeren Aufgabentext so schwierig mache, nannten Schüler zwei Gründe: Erstens müsse man als Schüler nicht nur verstehen, worum es in dem Text geht, sondern zweitens auch herausfinden, was man überhaupt machen soll. Problematisch sehen Schüler aber nicht nur die Aufgabenstellung, sondern auch die dazugehörige Antwort. Zwei Faktoren hatten hier einen besonderen Einfluss auf den Schwierigkeitsgrad. Zum einen war das der Komplexitätsgrad einer Antwort: je höher, desto schwieriger. Zum anderen waren da auch die Fachwörter, die den Schülern nicht geläufig waren. So antwortete eine Schülerin auf die Frage, warum gerade diese Aufgabe so schwierig sei: „weil mir die Wörter gefehlt haben, wie man es erklären soll“.

Neben den gerade aufgeführten schwierigkeitsgenerierenden Merkmalen nennen Schüler fehlendes **Vorwissen zum Thema (n=5)** und die **Bearbeitungszeit (n=3)**, die sie zur Lösung der Aufgabe benötigten. Hier wurde gesagt, dass eine Aufgabe umso schwieriger sein müsse, je mehr Zeit man für deren Bearbeitung benötigte.

6.8.4 Was eine Aufgabe nach Schülermeinung interessant macht

Schüler fanden anspruchsvollere Aufgaben interessanter als einfache (Tab. 36 auf Seite 155). Im vorherigen Abschnitt wurde bereits angedeutet, dass Schüler bei schwierigeren Aufgaben die Aufgabenstellung herausfordernder fanden, da zu deren Lösung **mehr als bloßes Entnehmen und Wiedergeben von Informationen (n=14)** nötig war. Fragt man Schüler nach einem Merkmal, das eine Aufgabe interessanter macht, so verwendeten sie zwar ähnliche Beschreibungen wie für schwierige Aufgaben, allerdings mit einer spielerischeren Konnotation. Es fielen Äußerungen wie: „da muss man grübeln... da muss man richtig kombinieren... da muss man sein Gehirn anstrengen... da muss man rumrätseln und dann will man es herausfinden... das ist Denksport!“

Die herausfordernde und spielerische Art einer Aufgabe ist allerdings nicht das wichtigste Kriterium einer Aufgabe. Das mit Abstand eindeutigste Merkmal für eine interessante Aufgabe aus Schülersicht ist der dabei erfahrene **Lernzuwachs (n=31)**. Hier-

zu äußerten Schüler, dass man sich bei einer interessanten Aufgabe wundert, staunt oder erst mal nicht glauben kann, dass es so etwas gibt, dass man nicht sofort eine Antwort darauf hat, sondern erst mal überlegen muss, um dann nach und nach zu einer Lösung zu kommen. Schüler beschrieben dies als ein Gefühl des Dazulernens, mehr zu behalten und schlauer zu werden. Einer der Schüler brachte es auf den Punkt: „Auf das Ergebnis komme ich nicht gleich und dann denke ich: Ah, jetzt weiß ich das auch (*grinst*)“.

Schüler beschrieben des Weiteren solchen Lernzuwachs als wertvoll, wenn sie den Eindruck hatten, dass ihnen die Lösungen Antworten auf Fragen in ihrem **Leben** oder ihrem **Alltag** boten. Sofern der Inhalt etwas mit der Lebenswelt der Schüler zu tun hatte, wurde die Kontextorientierung als interessanter empfunden (**n=11**). Ein annähernd gleichwertiger Faktor für Schülerinteresse war der Bezug der Aufgabe zum **eigenen Körper (n=7)**. Dies beinhaltete nach Schülermeinung sowohl das Erfahrbarmachen und Erlebarmachen, als auch das Erklären können von körpereigenen Funktionen und Prozessen: „weil es cool ist zu wissen“. Kurz nach dem eigenen Körper nannten Schüler **Tiere (n=6)** als interessante Inhalte von Aufgaben. In diesem Zusammenhang erwähnten sie auch, dass es interessanter ist, wenn eine Aufgabe in eine Geschichte oder einen **Kontext (n=4)** eingebunden ist als ohne. Auf der gleichen Stufe befindet sich **Kreativität (n=4)**. Dieses Merkmal wurde weder bei den Beschreibungen für leichte noch für schwere Aufgaben verwendet, sondern einzig zur Begründung von interessanten Aufgaben eingesetzt. Hierbei geht es um die Offenheit der Aufgabe und damit auch der Lösung. Am besten kann das durch die Metapher *thinking outside the box* beschrieben werden. Hierbei geht es hauptsächlich um unkonventionelles Denken. Für Aufgaben könnte das bedeuten, dass Aufgaben Gedankenexperimente zulassen sollten (was wäre wenn?) und nicht konkrete Lösungen geben oder einfordern sollten. Als eigenständiger Punkt wurde auch das Aufmerksam machen auf **Phänomene im Alltag (n=3)** genannt, die den Schülern bisher nicht bewusst aufgefallen waren (zum Beispiel dass nachts wirklich alle Katzen grau sind, solange sie nicht unter einer Laterne entlang laufen oder dass man bei Dämmerung mit den Augenwinkeln tatsächlich *besser* sehen kann). Hierbei beschrieben es die Schüler als besonders spannend, wenn sie dieses **Phänomen selbst wiederholen und ausprobieren (n=3)** konnten.

Abschließend werden in Tab. 38 die Einschätzungen der Schüler entsprechend dem Ranking dieser Studie absteigend aufgelistet.

Tab. 38: Ranking der Schüleraussagen zu den Aufgaben.

Was eine Aufgabe <u>einfach</u> macht	Was eine Aufgabe <u>schwierig</u> macht	Was eine Aufgabe <u>interessant</u> macht
Die mit der Aufgabe einhergehenden Lernmaterialien (n=21)	Die Anforderungen an die Schüler, die durch Aufgaben gestellt werden (n=16)	Der bei der Aufgabe erfahrene persönliche Lernzuwachs (n=31)
Die Reihenfolge der Aufgaben (n=5)	Das erforderliche Nachdenken (n=14)	Die Herausforderung durch die Aufgabenstellung (n=14)
Der Bekanntheitsgrad des Kontextes (n=3)	Die Masse an Aufgabentext und Inhalt (n=9)	Ein Zusammenhang mit der eigenen Lebenswelt (n=11)
Die Vorerfahrung der Schüler (n=3)	Das fehlende Vorwissen (n=5)	Der Bezug zum eigenen Körper (n=7)
	Die Bearbeitungszeit (n=3)	Der Bezug zu Tieren (n=6)
		Aufgaben, die in eine Geschichte oder einen Kontext gebettet sind (n=4)
		Das Zulassen von Kreativität (n=4)
		Das Beschreiben von Phänomenen im Alltag (n=3)
		Das selbst Wiederholen und Ausprobieren von Phänomenen im Alltag (n=3)

6.8.5 Qualität der Schülerantworten

In diesem Kapitel wird der Fokus auf die Qualität der Schülerantworten gelegt. Hierbei wurden, wie in den Kapitel davor, Schülergruppen gebildet und miteinander verglichen. Das methodische Vorgehen bei allen Vergleichen in diesem Kapitel erfolgte auf die gleiche Weise und wird im Folgenden exemplarisch für alle dargestellt.

Methodisches Vorgehen: Entsprechend der Einteilung der Schülerantworten in die Antwortkategorien *die Schülerantwort ist gut, schlecht oder gelingt nicht* (siehe Seite 169), konnten Aussagen zu den Unterschieden zwischen den Schülern in Bezug auf die Lernaufgaben getroffen werden. Um klare Aussagen über den Vergleich der Häufigkeit der Antwortkategorien (gut, schlecht, gelingt nicht) pro Schülergruppe zu machen, wurden die Antwortkategorien zusammengezählt und in ein Verhältnis zueinander gesetzt.

Abhängig davon, nach welchem Kriterium die Schüler in Gruppen eingeteilt wurden (Leistungsstärke, Geschlecht, Input...), unterschieden sich die Gruppen in Bezug auf die Anzahl der Personen pro Gruppe. Es gibt es zum Beispiel viel weniger geschlechts-homogene Dyaden als Schüler mit einem schlechten Ergebnis im Multiple-Choice-Test. Da sich die Schüleranzahl der Personengruppen zum Teil also erheblich unterscheiden, ist der Summenwert der jeweiligen Antwortqualität keine adequate Vergleichsgröße mehr.

Infolgedessen wurde der prozentuale Anteil der Antwortkategorie pro Gruppe gebildet und in den anschließenden Abbildungen dargestellt (siehe Abb. 37 bis Abb. 41). Deutliche Unterschiede in den Diagrammen wurden anschließend mit quantitativen Messmethoden untersucht. Da es sich bei den Messwerten um ordinal skalierte Daten handelt, wurde für die Überprüfung der Chi-Quadrat-Test (χ^2) verwendet. Statistisch signifikante Unterschiede wurden durch einen schwarzen Rahmen in den folgenden Balkendiagrammen kenntlich gemacht.

Ergebnisse des prozentualen Vergleichs der Qualität der Schülerantworten

1) **Fragestellung:** Wie häufig treten die Antwortkategorien *Die Antwort ist gut*; *die Antwort ist schlecht* und *die Antwort gelingt den Schülern nicht* bei den vier Lernaufgaben auf?

Ergebnis: Vergleicht man die prozentuale Häufigkeit der Antwortkategorien bei Aufgabe 1 fällt auf, dass 50% der Schüler eine gute Antwort, 33% eine schlechte Antwort geben und 17% der Schüler keine Antwort gelingt. Bei Aufgabe 2 steigert sich der Prozentsatz der guten Antworten auf 63% und der schlechten auf 32%. Nur 5% der Schüler findet keine Antwort. Dieses Ergebnis deckt sich mit der Einschätzung des Schwierigkeitsgrades der Aufgaben durch die Schüler (siehe Tab. 36 auf Seite 155).

Bei Aufgabe 3 geben 60% der Schüler eine gute Antwort, 20% eine schlechte Antwort und den restlichen 20% gelingt keine befriedigende Antwort. Die prozentuale Häufigkeit bei Aufgabe 4 ähnelt der Verteilung bei Aufgabe 3 sehr stark. Bei Aufgabe 4 geben 64% eine gute, 17% eine schlechte und 19% keine Antwort.

Vergleicht man die Verteilung der Schülerantworten über alle vier Aufgaben hinweg, so fällt auf, dass es bei den Aufgaben 1 und 2 zu einer graduellen Steigerung der Antwortqualität kommt, wohingegen die Verteilung bei Aufgabe 3 und 4 einer Stufung ähnelt. Beides wird durch die gestrichelten Linien in Abb. 37 verdeutlicht.

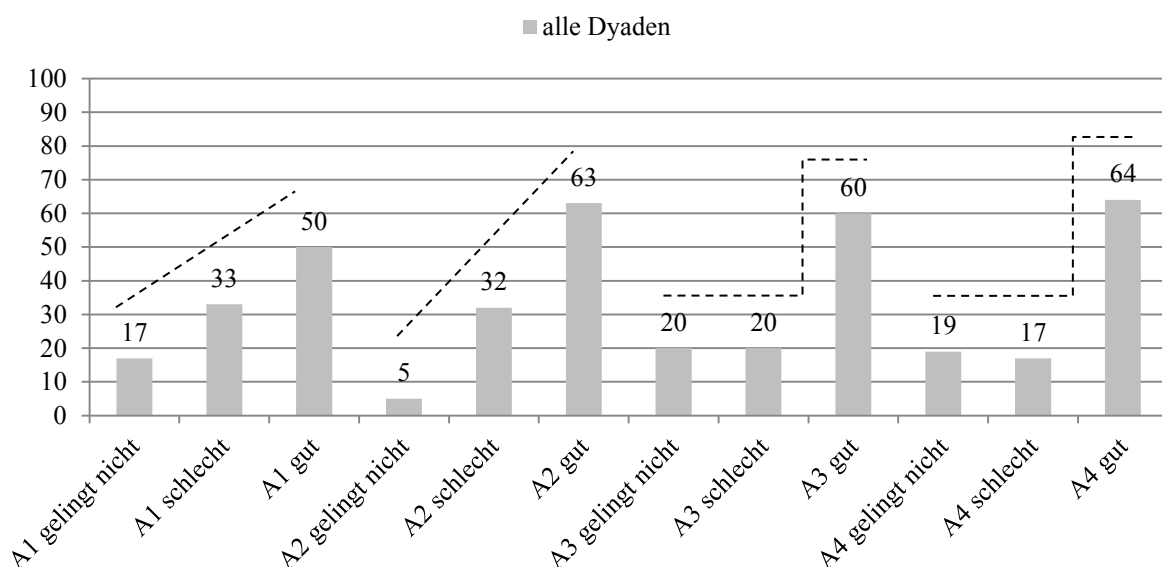


Abb. 37: Prozentuale Häufigkeit der Qualität der Schülerantworten bei den Aufgaben 1 bis 4.

Nach der Darstellung der prozentualen Häufigkeit der Antwortkategorien bei allen Schülern, folgt der Vergleich der Antworten von unterschiedlichen Schülergruppen. Wie bei allen anderen Gruppenvergleichen werden zuerst leistungsstärkere mit leistungsschwächeren Schülern verglichen, dann die Input-Gruppen, gefolgt von Mädchen und Jungen und dem Vergleich von Schülern, die ein gutes Ergebnis im Multiple-Choice-Test hatten, mit jenen, die ein schlechtes Ergebnis hatten.

2) **Fragestellung:** Wie häufig geben leistungsstärkere und leistungsschwächere Dyaden eine qualitativ gute, schlechte beziehungsweise keine Antwort?

Ergebnis: Es tritt im Prinzip das gleiche Muster wie bei Abb. 37 auf. Bei Aufgabe 1 und 2 kommt es zu einer „Steigung“, wohingegen sich bei Aufgabe 3 und 4 die Kategorien *Antwort gelingt nicht* und *Antwort ist schlecht* die Waage halten. Zur Verdeutlichung der Unterschiede im Muster der Daten, wurden auch in Abb. 38 gestrichelte Hilfslinien hinzugefügt.

Bei dem Vergleich der beiden Gruppen zeigte sich allerdings auch, dass leistungsstärkere Schüler (n=61) in der Regel qualitativ hochwertigere Antworten gaben und dass leistungsschwächere Schüler (n=49) besonders bei der Beantwortung von Aufgabe 1 Schwierigkeiten hatten (siehe Abb. 38). Dies konnte statistisch bestätigt werden. Leistungsschwächere Schüler erzielen bei Aufgabe 1 signifikant weniger gute Antworten als leistungsstärkere Schüler ($p = .003$).

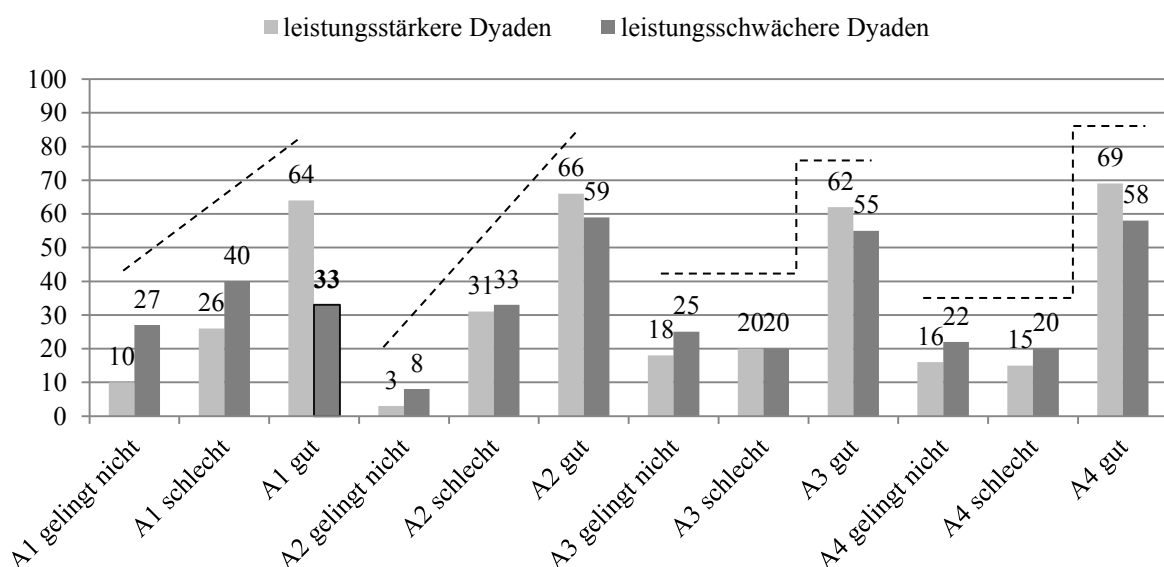


Abb. 38: Prozentuale Häufigkeit der Qualität der Schülerantworten bei den Aufgaben 1 bis 4 im Vergleich zwischen leistungsstärkeren und leistungsschwächeren Dyaden. Leistungsstärkere Dyaden erzielten bei allen Aufgaben prozentual eine gute Antwort.

3) **Fragestellung:** Wie häufig treten die Antwortkategorien bei Dyaden mit Text (n=55) im Vergleich zu Dyaden mit Experten-Concept-Map (n=55) auf?

Ergebnis: Der Input scheint keine nennenswerten Auswirkungen auf die Qualität der Antworten zu haben (siehe Abb. 39). Das deckt sich mit den Ergebnissen in Tab. 40 auf Seite 170. Bei der statistischen Überprüfung zeigte sich allerdings ein signifikanter Unterschied in der Qualität der Antworten bei Aufgabe 3 ($p = .035$).

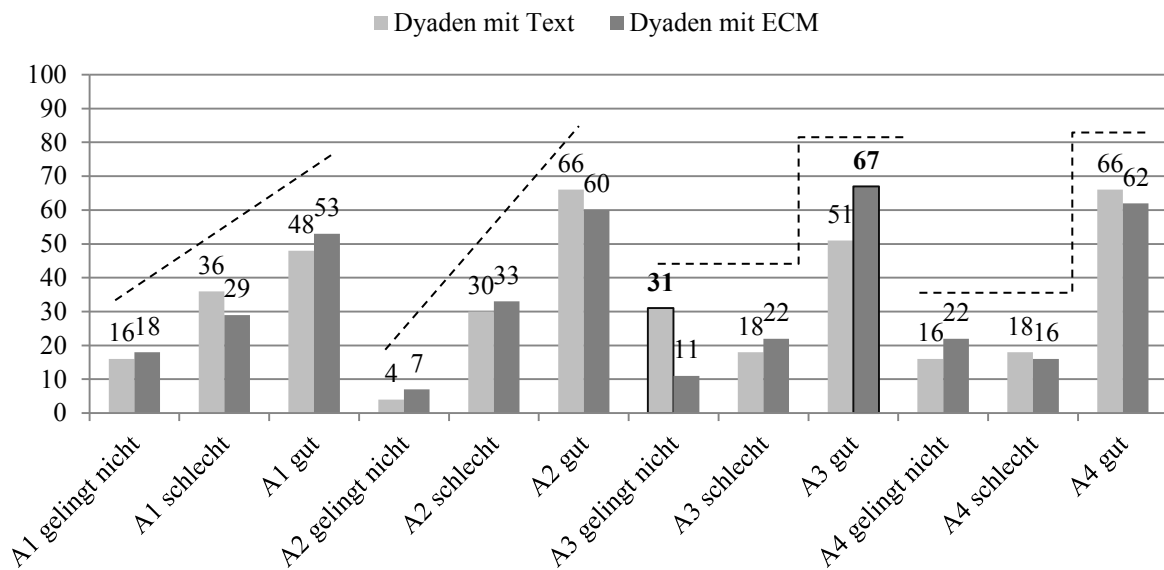


Abb. 39: Prozentuale Häufigkeit der Qualität der Schülerantworten bei den Aufgaben 1 bis 4 im Vergleich zwischen Dyaden mit Text und Dyaden mit Experten-Concept-Map als Input.

4) **Fragestellung:** Wie häufig treten die Antwortkategorien bei männlichen (n=28) im Vergleich zu weiblichen (n=40) Schülern auf?

Ergebnis: Bei der Aufschlüsselung der Daten in Bezug auf das Geschlecht tritt das bisherige Muster nicht mehr auf (vergleiche die Qualität der Schülerantworten bei den Mädchen in Abb. 40 mit Abb. 37 unter Berücksichtigung der gestrichelten Hilfslinien).

Besonders markant ist, dass die Qualität der Antworten bei den Aufgaben 3 und 4 bei den Mädchen im Vergleich zu den Jungen sichtbar schlechter ausfällt. Dieses Ergebnis konnte statistisch bestätigt werden. Die Qualität der Schülerantworten unterscheidet sich signifikant bei Aufgabe 3 und bei Aufgabe 4 (beide $p = .042$). Mädchen haben im Vergleich zu Jungen größere Schwierigkeiten die Aufgaben 3 und 4 zu beantworten.

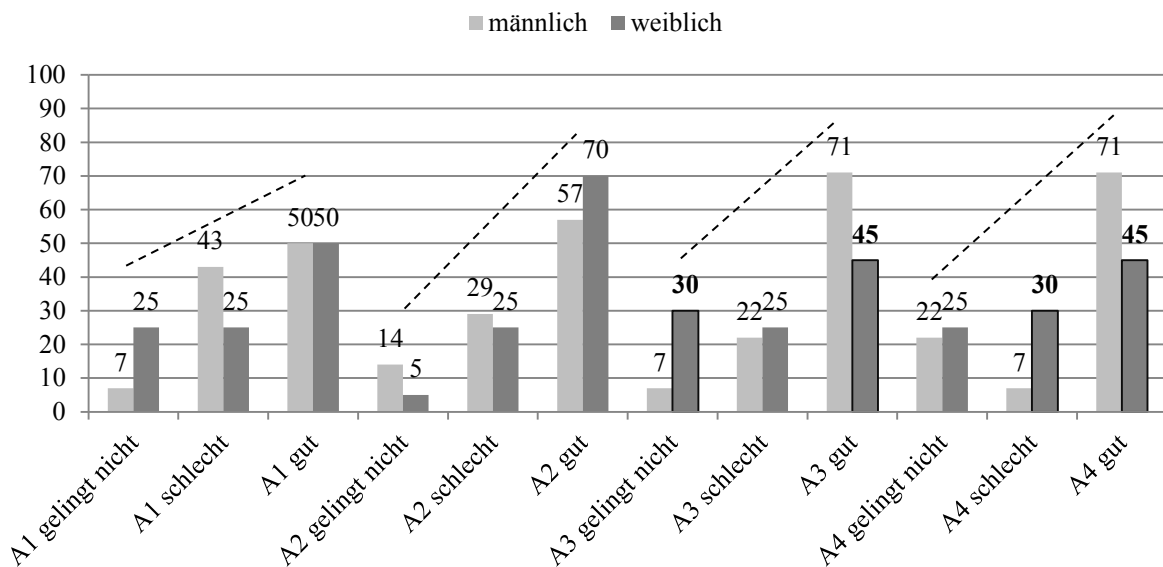


Abb. 40: Prozentuale Häufigkeit der Qualität der Schülerantworten bei geschlechtshomogenen Dyaden.

5) **Fragestellung:** Wie häufig treten die Antwortkategorien bei Schülern mit einem guten Ergebnis im Multiple-Choice-Test im Vergleich zu Schülern mit einem schlechten Ergebnis im Multiple-Choice-Test auf?

Ergebnis: Tendenziell fällt die Qualität der Antworten von Schülern mit einem guten Ergebnis im Multiple-Choice-Test besser aus, als bei Schülern mit einem schlechten Ergebnis. Dies zeigt sich besonders bei den Aufgaben 1 und 2. Bei der statistischen Überprüfung verfehlte das Ergebnis knapp das Signifikanzniveau von 5% und offenbarte einen statistisch marginal signifikanten Unterschied in der Qualität der Schülerantworten zwischen Schülern mit einem guten und Schülern mit einem schlechten Multiple-Choice-Ergebnis bei Aufgabe 1 ($p = .072$).

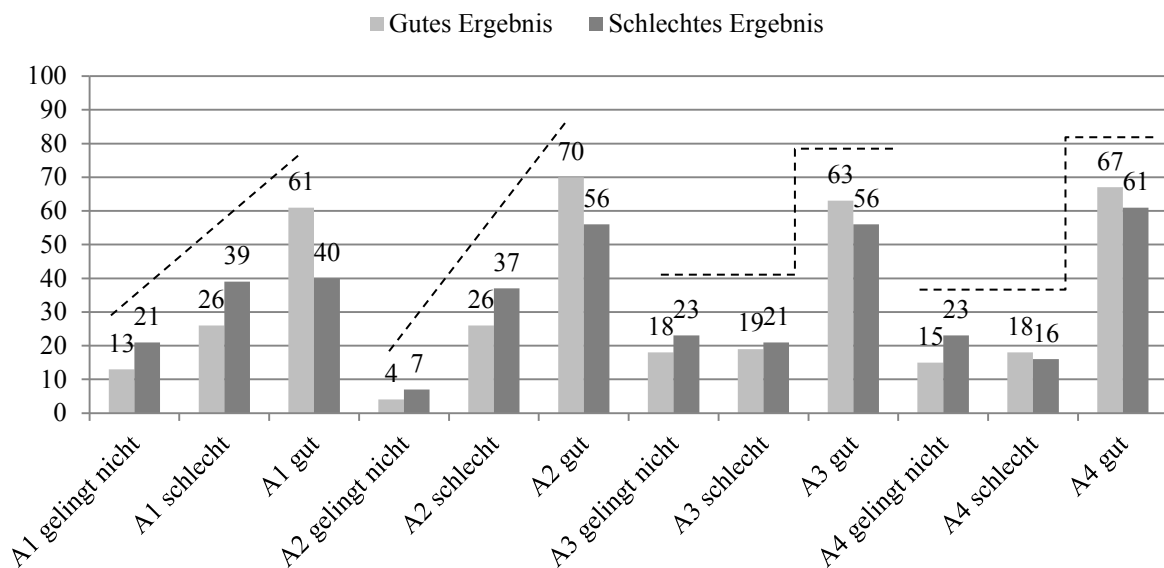


Abb. 41: Prozentuale Häufigkeit der Qualität der Schülerantworten bei Dyaden mit einem guten Ergebnis im Multiple-Choice-Test im Vergleich zu Dyaden mit einem schlechten Ergebnis im Multiple-Choice-Test.

6.8.6 Schwierigkeiten der Schüler bei der Interpretation der Abbildungen

Nach der Auseinandersetzung mit dem Input und dem Ausfüllen des Multiple-Choice-Tests erhielten die Schüler ein Arbeitsblatt mit Abbildungen zum menschlichen Auge (Arbeitsblatt siehe Seite 224). Das Arbeitsblatt setzte sich aus insgesamt 5 Abbildungen zusammen. Die 5 Abbildungen standen untereinander im Zusammenhang und folgten der gleichen farblichen Kodierung wie der Input. Bei Abbildung 3 war exemplarisch ein Abbildungstitel aufgeführt, der bei allen andern Abbildungen fehlte.

6.8.6.1 Abbildung 4 ist für alle Schüler am schwierigsten

Fragestellung: Bei welcher Abbildung hatten Schüler die meisten Probleme?

Methodisches Vorgehen: Die Schüler wurden aufgefordert, sich nacheinander die Abbildungen anzuschauen und dann mögliche Titel für jede Abbildung zu finden. Um es den Schülern leichter zu machen, erhielten sie ein weiteres Arbeitsblatt mit vorformulierten Titeln (siehe Seite 224). Hier wurden pro Abbildung immer fünf mögliche Titel aufgeführt. Von den fünf Titeln waren immer zwei richtig, allerdings war dies den Schülern bei der Bearbeitung nicht bekannt. Genauso wie beim Multiple-Choice-Test wurden richtig angekreuzte und richtig nicht angekreuzte Items mit jeweils einem Punkt versehen und miteinander verrechnet. Die daraus resultierende theoretische Punktzahl reichte von 0 bis 20.

Ergebnis: Die von den Schülern tatsächlich erreichte Gesamtpunktzahl für alle Abbildungen reicht von 10 bis 17. Bei einer Spannweite von 7 Punkten lag der Mittelwert bei 14,06 und der Median bei 14. Die meisten Punkte wurden bei Abbildung 5 und die wenigsten bei Abbildung 4 erzielt. Dies bedeutet, dass den Schülern Abbildung 5 am wenigsten und Abbildung 4 am meisten Probleme gemacht hat (siehe Tab. 39).

Tab. 39: Punkteanzahl im Multiple-Choice-Test zu den Abbildungen.

	Anzahl der Punkte für Abbildung 1	Anzahl der Punkte für Abbildung 2	Anzahl der Punkte für Abbildung 4	Anzahl der Punkte für Abbildung 5	Gesamtpunktzahl im Multiple-Choice-Test
N	112	112	112	112	112
Mittelwert	3.62	3.64	2.82	3.98	14.06
Median	4.00	4.00	2.00	4.00	14.00
Minimum	1	2	1	2	10
Maximum	5	5	5	5	17
Summe	405	408	316	446	1575

Vergleicht man die Abbildungen untereinander, ist das Ergebnis nicht verwunderlich. Abbildung 4 entspricht der höchsten Vergrößerungs- und Abstraktionsstufe und ist somit am weitesten von der (be-)greifbaren Welt der Schüler entfernt. Des Weiteren enthält Abbildung 4 die meisten neu erlernten Strukturen und beschreibt zudem einen besonderen Fall im Aufbau der Netzhaut, die den Schülern ohnehin am meisten Schwierigkeiten bereitete (vergleiche Seite 144 ff.). Um bei ähnlichen Aufgaben die Schwierigkeiten bei der Entnahme von Informationen aus Abbildungen vorzubeugen, sollte darauf geachtet werden, das Abstraktionsniveau niedrig zu halten und die Informationsdichte zu verringern.

6.8.6.2 Leistungsschwächere haben Schwierigkeiten mit Abbildungen

Fragestellung: Gibt es bei den verschiedenen Schülergruppen Unterschiede bei dem Verstehen der Abbildungen? (Die Abbildungen befinden sich auf Seite 224)

Methodisches Vorgehen: Um Unterschiede im Abbildungsverstehen der Schüler zu finden, wurden die Mittelwerte in weiteren t-Tests in Bezug auf die Untergruppen dieser Studie verglichen (gleiches methodisches Vorgehen wie bei Punkt 6.8.6.1).

Ergebnisse:

- Es konnten keine statistisch signifikanten Unterschiede zwischen leistungsstärkeren ($n=51$) und leistungsschwächeren Schülern ($n=61$) gefunden werden. Bei Abbildung 2 trat allerdings ein marginal signifikanter Unterschied zu Gunsten von leistungsstärkeren Schülern auf ($p = .074$).
- Es konnte kein statistisch signifikanter Unterscheid zwischen Schülern mit Text ($n=57$) und Schülern mit Experten-Concept-Maps ($n=55$) in Bezug zum Abbildungsverständnis gefunden werden.
- Ferner konnte kein statistisch signifikanter Unterscheid zwischen Mädchen ($n=62$) und Jungen ($n=50$) gefunden werden.
- Bei Abbildung 5 konnte ein Unterschied zwischen Schülern mit einem guten Ergebnis ($n=56$) im Multiple-Choice-Test und Schülern mit einem schlechten Ergebnis ($n=56$) gefunden werden. Schüler mit einem guten Ergebnis im Wissenstest erzielten statistisch signifikant mehr Punkte bei Abbildung 5 als Schüler mit einem schlechten Ergebnis im Wissenstest ($p = .046$).

6.8.7 Einfluss des Inputs auf die Antwortqualität bei Reproduktionsaufgaben

Sofern Schüler aus einem der beiden Formen des Inputs mehr Informationen entnommen haben, könnte man davon ausgehen, dass der bessere Input auch zu einer Steigerung der Qualität der Schülerantworten führt. Hieraus ließ sich folgende Fragestellung ableiten:

Fragestellung: Inwieweit gibt es einen Zusammenhang zwischen dem Multiple-Choice-Ergebnis und der Qualität der Schülerantworten bei den Lernaufgaben 1 – 4?

Beschreibung der Vorgehensweise: Die Lernaufgaben (siehe Seite 112) wurden den Schülern nacheinander im Anschluss an den Multiple-Choice-Test gegeben. Im Gegensatz zu den schriftlichen Multiple-Choice-Ergebnissen sollten sich die Schüler bei den Lernaufgaben ausdrücklich unterhalten. In Partnerarbeit sollten sie gemeinsam eine Antwort finden und diese dann dem Versuchsleiter mündlich mitteilen.

Methodisches Vorgehen: Bei der anschließenden Videoanalyse der Interviews wurden die Schülerantworten pro Aufgabe einer von drei Kategorien zugeordnet. Die drei Kategorien waren: die Antwort gelingt den Schülern nicht, die Qualität der Antwort ist schlecht, die Qualität der Antwort ist gut (Interrater-Reliabilität und Kodierungsregeln siehe Seite 118).

Auswertung: Für die quantitative Auswertung wurden die Schülerantworten entsprechend ihrer Qualität mit Punkten versehen (*Antwort gelingt nicht* = 0 Punkte; *Antwort ist schlecht* = 1 Punkt; *Antwort ist gut* = 2 Punkte). Die Punkte pro Aufgabe wurden addiert und ergaben das Gesamtergebnis für die Qualität der Schülerantworten über alle vier Aufgaben hinweg.

Ergebnis: Das daraus resultierende Spektrum ist intervallskaliert und reicht von 0 bis 8 Punkten. Die Spannweite liegt bei 7 und der Mittelwert bei 5,73. Der Median ist 6,00 und die Standardabweichung 2,054. Um die Frage zu beantworten, ob das Multiple-Choice-Ergebnis in Zusammenhang mit der Qualität der Schülerantworten stand, wurde eine Korrelation nach Pearson gerechnet, die das Multiple-Choice-Ergebnis aller Schüler mit deren Ergebnis bei den Lernaufgaben vergleicht. Die Berechnungen wurden für alle Untergruppen wiederholt.

Die Ergebnisse in Tab. 40 zeigen keine signifikanten Werte, das heißt bei allen Gruppen ging das Mehr an Wissen durch den Input nicht messbar mit besseren Antworten bei Aufgaben der Reorganisation und Anwendung einher. Es konnte also nicht gezeigt werden, dass sich durch die Form des Inputs das Verstehen verbessert.

Tab. 40: Korrelationen zwischen dem Wissenstest und der Qualität der Schülerantworten für die Aufgaben 1 bis 4 bei unterschiedlichen Schülergruppen. Bei keiner Schülergruppe traten signifikante Korrelationen nach Pearson zwischen den Multiple-Choice-Ergebnissen und der Qualität der Schülerantworten bei den Lernaufgaben auf.

		Punktezahl im MC-Test (ohne Aufgabe 1)	Gesamtqualität der Schülerantwort bei allen Aufgaben zusammen
Alle Schüler	Korrelation nach Pearson	.082	.082
	5%-Signifikanzniveau (2-seitig)	± .404	±.404
	N	106	106
Leistungsstärkere Schüler	Korrelation nach Pearson	.150	.150
	5%-Signifikanzniveau (2-seitig)	± .258	±.258
	N	59	59
Leistungsschwächere Schüler	Korrelation nach Pearson	.022	.022
	5%-Signifikanzniveau (2-seitig)	± .884	±.884
	N	47	47
Text	Korrelation nach Pearson	.235	.235
	5%-Signifikanzniveau (2-seitig)	± .094	±.094
	N	52	52
Experten-Concept- Map	Korrelation nach Pearson	-.056	-.056
	5%-Signifikanzniveau (2-seitig)	± .685	±.685
	N	54	54
Männlich	Korrelation nach Pearson	.273	.273
	5%-Signifikanzniveau (2-seitig)	± .057	±.057
	N	49	49
Weiblich	Korrelation nach Pearson	-.004	-.004
	5%-Signifikanzniveau (2-seitig)	± .974	±.974
	N	61	61
Schüler mit einem guten MC-Ergebnis	Korrelation nach Pearson	-.166	-.166
	5%-Signifikanzniveau (2-seitig)	± .231	±.231
	N	54	54
Schüler mit einem schlechten MC- Ergebnis	Korrelation nach Pearson	-.007	-.007
	5%-Signifikanzniveau (2-seitig)	± .957	±.957
	N	56	56

6.9 Zusammenfassung

Im Folgenden werden die Einzelergebnisse zusammengefasst und in Zusammenhang zueinander gesetzt. Zur besseren Orientierung bei den Ergebnissen wird in Abb. 42 nochmals der Ablauf der Interviews dargestellt sowie die Aufgabentexte mit den Schwierigkeitsgraden aufgeführt. Die kognitiven Operationen, die für das Lösen der Aufgaben notwendig sind, werden in den Hilfen genannt (siehe Seite 235 ff.).

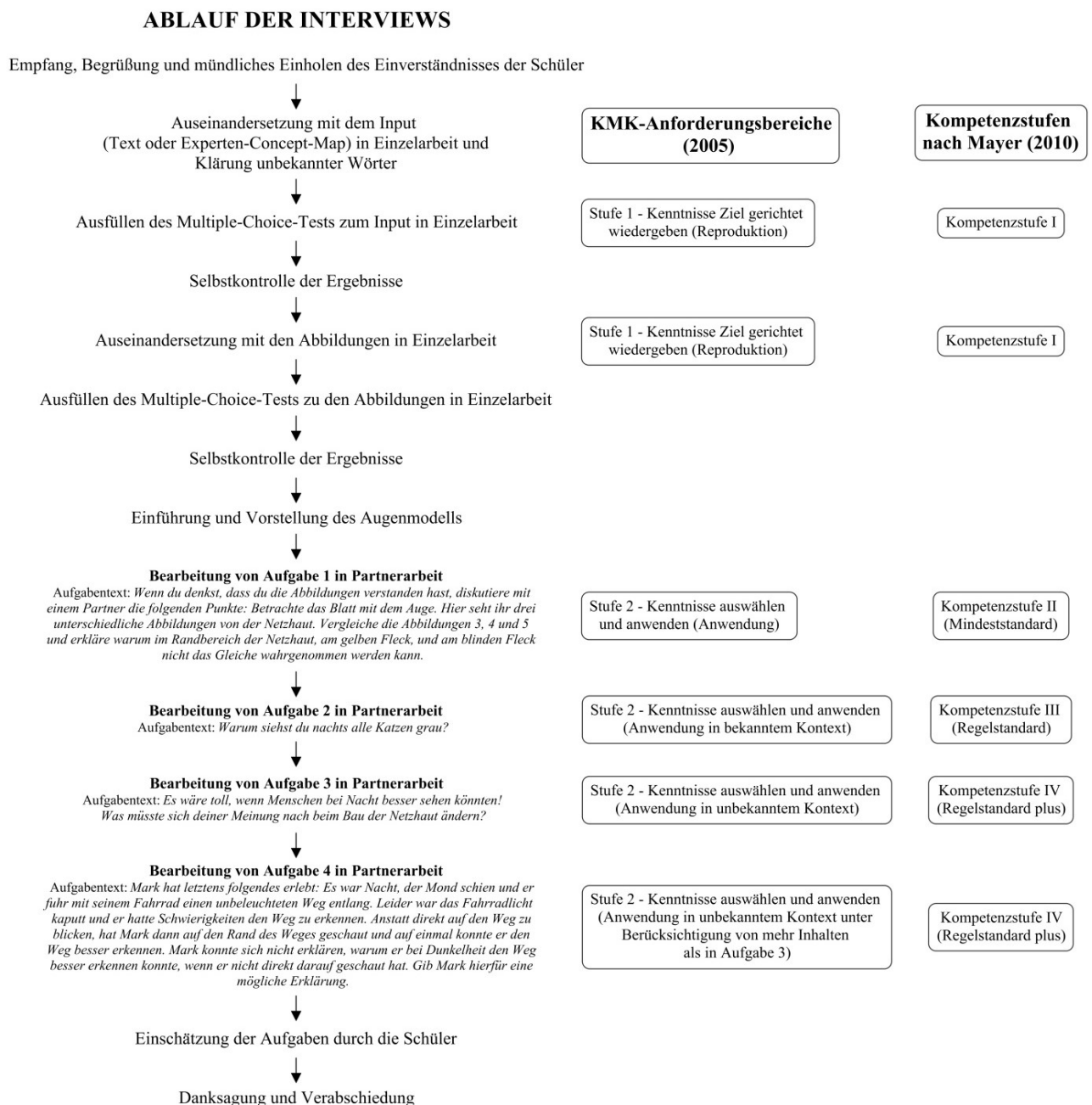


Abb. 42: Ablauf der Interviews mit Aufgabentexten (siehe Seite 112) und Schwierigkeitsgraden (siehe Seite 22) entsprechend der Anforderungsbereiche der KMK (2004) und dem Kompetenzstufenmodell nach Mayer (2013).

- 1) **Die Schüler zeigten individuelle Lernwege.** Bei den Lernaufgaben wurde schon bei der Erprobung der Materialien klar, dass das Problemlöseverhalten der Schüler nicht identisch verläuft. Dies zeigte sich daran, dass Schüler durch die Impulse der Hilfen auf andere Dinge hingewiesen wurden, als solche, über die sie gerade diskutierten. Anders gesagt, die Schüler sind in Teilen unterschiedlich an die Lösung der Aufgaben herangegangen als dies im Vorfeld geplant war (siehe Seite 126).
- 2) **Das Augenmodell fördert Verstehen und Kommunikation.** Neben gestuften Hilfen, wurde in dieser Studie eine Verbalisierungshilfe mittels eines Augenmodells angeboten. Dieses Augenmodell förderte das Verstehen und die Kommunikation über Inhalte. Bei dessen Einsatz wurde allerdings auch deutlich, dass für viele Schüler der Umgang mit dem Modell nicht vertraut war (siehe 6.2.2 auf Seite 126).
- 3) **Alle Schüler profitierten bei Reproduktionsaufgaben von Experten-Concept-Maps.** Die Ergebnisse dieser Studie haben gezeigt, dass sich der Einsatz von Experten-Concept-Maps positiv auf das Wissen aller Schüler im Bereich der Reproduktion auswirkt. In Bezug auf den Input zeigte sich kein Unterschied zwischen Mädchen und Jungen. Jedoch erzielten leistungsschwächere Schüler mit einer Experten-Concept-Map marginal signifikant bessere Ergebnisse als leistungsschwächere Schüler mit einem Fachtext (siehe Seite 127 ff.).
- 4) **Mädchen hatten zwar bessere Noten, zeigten aber keine besseren Ergebnisse.** Mädchen hatten in den Fächern Biologie und Deutsch signifikant bessere Noten als Jungen (außer in Mathematik), allerdings passte das nicht zu deren erbrachten Leistungen in dieser Studie. Besonders markant waren die geschlechtlichen Unterschiede bei den Aufgaben 3 und 4. Hier erzielten Mädchen signifikant schlechtere Antworten und benötigen mehr Zeit und Hilfe als Jungen. Bei der Untersuchung der Art von Hilfe zeigte sich, dass Mädchen besonders beim Lichteinfall ins Auge und dem Aufbau der Netzhaut Schwierigkeiten hatten (siehe Seite 129 ff.).
- 5) **Der Zeitbedarf der Schüler war unterschiedlich.** Bei leistungsschwächeren Schülern konnte ein großer Einfluss der Bearbeitungszeit des Inputs auf das Ergebnis des MC-Tests festgestellt werden. Es hat sich gezeigt, dass das Ergebnis im Multiple-Choice-Test nachweisbar besser wurde, je mehr Zeit sich Schüler für die Erarbeitung genommen haben (siehe Seite 132 ff.).
- 6) **Leistungsschwächere Schüler brauchen mehr Betreuung.** Am Beispiel von Aufgabe 4 konnte gezeigt werden, dass sowohl leistungsstärkere als auch leistungsschwächere Schüler im Durchschnitt mit annähernd der gleichen Sprechzeit zur Lösung kamen. Eine genaue Untersuchung der Sprechzeit ergab, dass leistungsstärkere Gruppen einen höheren Redeanteil im Vergleich zu leistungsschwächeren Gruppen bei Aufgabe 4 im Interview hatten. Es hat sich allerdings auch gezeigt, dass sich die Sprechzeit des Versuchsleiters ändert. Bei leistungsstärkeren Gruppen sprechen die Schüler etwas mehr und der Versuchsleiter weniger, bei leistungs-

schwächeren Gruppen sprechen die Schüler etwas weniger, dafür aber der Versuchsleiter mehr (siehe Seite 133 ff.).

- 7) **Gestufte Hilfen bilden fachliche und nicht-fachliche Verstehensschwierigkeiten von Schülern ab und machen Unterschiede zwischen Schülergruppen deutlich.** Das individuelle (nicht lineare) Geben von gestuften Hilfen hat gezeigt, dass es zu Häufungen bestimmter Hilfen bei den Aufgaben kam (siehe Tab. 33 auf Seite 145). Die Häufungen erlaubten durch eine Kategorisierung der Hilfen ein Ranking der Verstehensschwierigkeiten der Schüler bei den Aufgaben (siehe Kapitel 6.7 ab Seite 146). Bei den Vergleichen der Verstehensschwierigkeiten zwischen unterschiedlichen Schülergruppen hat sich gezeigt, dass leistungsschwächere Schüler bei Aufgabe 1 in Bezug auf das Abbildungsverstehen Schwierigkeiten hatten (siehe Abb. 33 auf Seite 150) und Mädchen bei den Aufgaben 3 und 4 mehr Hilfe als Jungen benötigten (siehe Abb. 35 auf Seite 152).
- 8) **Ein Viertel aller Hilfen gingen über rein fachliche Unterstützung hinaus.** Unabhängig von der Gruppenzugehörigkeit oder dem Geschlecht benötigte mindestens ein Viertel aller Schüler nicht-fachliche Unterstützung beim Lösen der Lernaufgaben (siehe Kapitel 6.6.2 auf Seite 145).
- 9) **Leistungsschwächere Schüler benötigten mehr Hilfe bei der Erschließung von Inhalten.** Generell benötigten leistungsschwächere Schüler bei allen Aufgaben mehr Unterstützung als leistungsstärkere Schüler. Markante Unterschiede traten allerdings bei Aufgabe 1 auf. Hier erzielten leistungsschwächere Schüler signifikant schlechtere Antworten, benötigen mehr Zeit und Hilfe. Diese Hilfen standen primär in Zusammenhang mit genauem Betrachten der Abbildungen und der Formulierung der Antwort (siehe Seite 148 ff.).
- 10) **Die geplanten Schwierigkeitsgrade bei der Aufgabenkonstruktion wurden von den Schülern bestätigt.** Die deskriptive Beschreibung der Aufgaben hat in Verbindung mit den qualitativen Daten (Lösungshäufigkeit, Qualität der Schülerantworten, Anzahl der Hilfen) gezeigt, dass der Schwierigkeitsgrad einer Aufgabe im Vorfeld zu Teilen planbar ist (siehe Seite 144 ff.). Darüber hinaus hat die Einschätzung der Aufgaben durch die Schüler gezeigt, dass Schülern die Schwierigkeit von Aufgaben bewusst ist. Es hat sich allerdings auch gezeigt, dass der Schwierigkeitsgrad einer Aufgabe individuell unterschiedlich empfunden wird (siehe hierzu Seite 154 ff.).
- 11) **Schwierige Aufgaben waren für Schüler interessanter.** Des Weiteren waren die Beschreibungen der Schüler, was eine Aufgabe schwierig oder interessant macht, zu einem großen Teil deckungsgleich. Das machte die Trennung der beiden Kategorien nicht einfach. Fakt bleibt, dass alle Schüler schwierigere Aufgaben interessanter empfanden, da sie durch diese Aufgaben den größten Lernzuwachs bei sich selbst feststellen konnten (siehe Seite 158 ff.).

- 12) **Schüler, die schon mit dem Verstehen des Inputs Schwierigkeiten hatten, hatten auch Probleme beim Verstehen der Abbildungen und beim Lösen der Aufgaben.** Bei den Vergleichen von Schülern, die über ihr Multiple-Choice-Ergebnis in Gruppen geteilt wurden, traten des Weiteren Gemeinsamkeiten zu jenen Schülern auf, die über ihre Biologienote in Gruppen geteilt wurden. Genauso wie bei den leistungsschwächeren Schülern schnitten Schüler mit einem schlechteren Ergebnis im Multiple-Choice-Test bei Aufgabe 1 weniger erfolgreich ab als leistungsstärkere Schüler oder Schüler mit einem guten Ergebnis im Multiple-Choice-Test. Unabhängig von den Lernaufgaben und dem Input erzielten leistungsschwächere Schüler und Schüler mit einem schlechteren Ergebnis im Multiple-Choice-Test darüber hinaus weniger Punkte im Verstehens-Test der Abbildung als leistungsstärkere Schüler und Schüler mit einem besseren Ergebnis im Multiple-Choice-Test (siehe Seite 167 ff.).
- 13) **Die Form des Inputs hatte keinen Einfluss auf die Antwortqualität der Schüler bei Transfer- und Anwendungsaufgaben.** Der nachgewiesene Wissensvorsprung durch Experten-Concept-Maps bei Aufgaben der Reproduktion hatte über alle Schüler hinweg keinen messbaren Einfluss auf die darauffolgenden Lernaufgaben. So hatte der unterschiedliche Input weder Auswirkungen auf die Qualität der Schülerantworten bei Aufgaben der Reorganisation und Transfer, noch auf die Anzahl oder die Art der benötigten Hilfen. Der Input machte auch keinen zeitlichen Unterschied bei der Beantwortung der Lernaufgaben aus (siehe Seite 169 ff.).

7 DISKUSSION

Dem in dieser Forschungsarbeit entwickelten Differenzierungsansatz ging eine Befragung von Biologielehrkräften voraus. Es sollte geklärt werden, welche Art von innerer Differenzierung sie sich in ihrem Unterricht wünschen würden. Die Resultate dieser Befragung bildeten die Grundlage für den Differenzierungsansatz, der in Kooperation mit Biologie-Lehrkräften entwickelt und empirisch überprüft wurde (siehe die Zusammenfassung in Kap 4.4 auf Seite 103).

Das Ergebnis der Überprüfung zeigte, dass Experten-Concept-Maps den Wissenserwerb während individueller Lernphasen bei Schülern fördern. Weiter konnte durch eine explorative Datenanalyse gezeigt werden, dass die Bearbeitungszeit bei leistungsschwächeren Schülern eine Auswirkung auf das Ergebnis hat und je nach Schüler unterschiedliche Arten von Hilfen beansprucht werden. Es offenbarte sich auch, dass sich die Schulnoten der Schüler in diesem Projekt nicht mit der von ihnen erbrachten Leistung in dieser Studie in Einklang bringen ließen. Eine Zusammenstellung der Ergebnisse findet sich unter 6.9 ab Seite 171.

In der folgenden Diskussion werden die markantesten Ergebnisse aufgegriffen und diskutiert. Sofern sich daraus Empfehlungen für die Schulpraxis ableiten lassen, so werden diese zu den jeweiligen Punkten formuliert.

7.1 Zu der Lehrerfortbildung

Lehrkräfte benötigen alltagstaugliche Differenzierungsmethoden

Obwohl bereits eine Vielzahl an Möglichkeiten zur inneren Differenzierung existiert (siehe Abb. 6 auf Seite 22), wurde in unserer Lehrerfortbildung deutlich, dass nur wenige davon als alltagstauglich für die Schulpraxis erachtet werden (siehe 4.3 ab Seite 97). Die Ergebnisse sind aufgrund der geringen Anzahl an teilnehmenden Lehrkräften zwar nicht repräsentativ, können aber dennoch die Grundlage für die Entwicklung eines Fragebogens bilden. Dieser Fragebogen könnte sowohl den Bedarf von Methoden zur Binnendifferenzierung (und Inklusion!) für Biologielehrkräften ermitteln, als auch die Voraussetzung für das Gelingen von Differenzierung im Biologieunterricht aus Sicht der Lehrer an einer repräsentativen Anzahl an Biologielehrkräften abfragen.

Nichtsdestotrotz entsprechen die qualitativen Ergebnisse aus der Lehrerfortbildung den Ergebnissen internationaler Vergleichsstudien wie zum Beispiel PISA (Klieme et al., 2010). Die PISA-Studien (2000, 2003 und 2006) haben gezeigt, dass leistungsschwächere Schüler und besonders Schüler mit Migrationshintergrund in ihrer Scientific

Literacy stark zurück liegen (Stanat et al., 2010). Daher muss diesen Schülern adäquate Unterstützung zukommen, was im Biologieunterricht allerdings bisher kaum der Fall ist (Jatzwauk und Sandmann 2007).

Für zukünftige fachdidaktische Forschung, in Bezug auf Binnendifferenzierung und Leistungsheterogenität, muss erwähnt werden, dass eine nachgewiesene Wirkung einer Intervention nicht automatisch zu deren Umsetzung in der Schule führt. Nach Aussage der befragten Lehrer werden bei vielen Maßnahmen nicht alle Facetten des Schulalltages berücksichtigt. Dies umfasst die Unterrichtsplanung sowie die praktische Umsetzung bis hin zur Notengebung. Als Empfehlung zur Steigerung der Praxistauglichkeit von neuen Methoden sollten von Beginn an mehr Lehrkräfte am Entwicklungs- und Forschungsprozess beteiligt werden. Durch solch eine Partizipation könnten, gerade bei der Neuentwicklung von Methoden, frühzeitig wichtige Impulse mit eingebracht werden. Die partizipative Methodenentwicklung würde zu dem die Chancen erhöhen, dass die neuen Methoden besser von anderen Lehrern angenommen werden und somit mehr Binnendifferenzierung im Biologieunterricht stattfinden könnte.

7.2 Zu der formativen Evaluation der Materialien

7.2.1 Schüler zeigten unterschiedliche Lernwege

Bei der Durchführung der Interviews hat sich gezeigt, dass die Bearbeitung der Hilfen in strenger Reihenfolge (1 bis 5) in dem Setting einer Interviewsituation unvorteilhaft war (siehe 6.2.1 auf Seite 126). So kam es vielmehr zu Behinderungen bei der Auseinandersetzung mit den Aufgaben, als dass der Verstehens-Prozess erleichtert worden wäre.

Dass Schüler unterschiedlich an ein Problem oder eine Aufgabe herangehen, ist nicht verwunderlich. Bereits bei Grundschulkindern konnten durch qualitative Untersuchungen verschiedene Lösungsstrategien beim Bearbeiten von Mathematikaufgaben gefunden werden (Heinze 2005). So scheinen Grundschüler ungleiche Problemlösestile zu besitzen, die sich in einer unterschiedlichen Herangehensweise an Probleme verdeutlichen (Fuchs 2006). Demzufolge ist es sehr wahrscheinlich, dass auch bei Lernaufgaben in der Biologie verschiedene Problemlösetypen existieren (vergleiche Lind et al. 2004 in Biologie und Günther 2013 in Mathematik).

Die Arbeiten von Friege (2001) und Czeskleba (2012) deuten darauf hin, dass Schüler keine allgemeine, sondern nur eine domänenspezifische Problemlösefähigkeit besitzen, die auf Praxis, Training und Erfahrung beruht. Aus diesem Grund kann die Frage nach individuellen Lernwegen auch auf die Biologie übertragen und entsprechend untersucht werden. Ließen sich nämlich unterschiedliche Problemlösetypen

beim Bearbeiten von Aufgaben mit biologischen Fachinhalten (Methoden und/oder Repräsentationsformaten) identifizieren, so wären diese ein weiterer Ansatz für eine innere Differenzierung im Biologieunterricht.

7.2.2 Modelle fördern Verstehen

Bei der formativen Evaluation der Materialien hat sich ferner gezeigt, dass viele Schüler Probleme dabei hatten, sich den Einfall des Lichtes in das Auge und auf die Retina vorzustellen (siehe 6.2.2 auf Seite 126). Zur Förderung der Kommunikation und der Anschaulichkeit wurde zusätzlich ein Modell des menschlichen Auges eingesetzt. Dieses Modell half den Schülern, ihre Gedanken trotz fehlender Worte auszudrücken. Durch dessen Einsatz und den Änderung der Durchführung in Bezug auf die Reihenfolge der Hilfen (siehe 7.2.1 ab Seite 176) haben Schüler entweder fachlich richtig geantwortet oder es konnte die richtige Lösung mit den Schülern besprochen werden. Die Anschaulichkeit des Modells erleichterte den Schülern die Aufnahme und das Einprägen von Strukturen. Das eingesetzte Modell unterstützte die Lernenden bei ihrem Erkenntnisprozess darüber hinaus durch seine Reduktion, indem es den Blick auf das Wesentliche lenkte. Ferner lieferte es einen Blickwinkel auf Strukturen und Zusammenhänge, die in der Realität kaum anschaulich waren (Köhler 2010, S. 168). Die Lernwirksamkeit von Modellen im Biologieunterricht ist anerkannt und wird an dieser Stelle nicht weiter diskutiert. Weiterführende Literatur siehe: Leibold und Klautke 1999, Terzer und Upmeier zu Belzen 2007, Upmeier zu Belzen und Krüger 2010, Meisert 2012.

7.3 Zu den Ergebnissen von Experten-Concept-Maps und Text

7.3.1 Alle Schüler profitieren von Experten-Concept-Maps bei reiner Reproduktion

Während die Schüler den Input bearbeiteten und den Multiple-Choice-Test ausfüllten, ging es um das Entnehmen und Reproduzieren von Informationen. Hierbei mussten die Schüler keine Informationen selektieren, organisieren oder integrieren (Kauertz et al. 2010), sondern einfach nur eine Information erinnern.

Es hat sich gezeigt, dass Lernende mit einer Experten-Concept-Map bei Reproduktionsaufgaben bessere Ergebnisse erzielen als mit einem Fachtext (siehe 6.3.1 auf Seite 127). Dieses Resultat unterstützt die Ergebnisse von Hauser (2006). Das Ergebnis dieser Studie zeigt aber auch, dass Experten-Concept-Maps nicht zu einem tieferen Verständnis der Inhalte oder verbessertem Transfer führen. Der einmalige Einsatz von Experten-Concept-Maps scheint hierfür nicht auszureichend zu sein. Um Transferleistungen zu steigern, benötigt es neben dem entsprechenden Fachwissen vor allem

mehr Übung (vgl. Friege 2001, Lind et al. 2004). Es konnten keine Effekte zwischen dem Input und der Qualität der Schülerantworten bei Lernaufgaben gefunden werden, die über die Reproduktion von Wissen hinausgingen (siehe 6.3.1 ab Seite 127). Dies scheint nicht verwunderlich, da Schüler mit dem Format der Reorganisations- und Transferaufgabe wenig vertraut sind (Haugwitz 2009).

Genauso wie Abituraufgaben weniger auf komplexe kognitive Prozesse abzielen sondern eher kleinschrittige Handlungsanweisungen mit reproduktivem Charakter darstellen (Kühn 2011), verhält es sich im Unterricht. Anhand einer Videoanalyse konnte nachgewiesen werden, dass im Biologieunterricht signifikant mehr Aufgaben mit geringem kognitiven Anspruch gestellt werden (Jatzwauk et al. 2008).

Als Empfehlung lässt sich hieraus formulieren, dass Reorganisation und Transfer im Unterricht stärker geübt und mehr Aufgaben mit einem höheren kognitiven Anspruch gestellt werden sollten (Mandl et al. 2004, Klauer 2010).

7.3.2 Leistungsschwächere Schüler profitieren marginal von Experten-Concept-Maps

Ferner konnte gezeigt werden, dass leistungsschwächere Schüler marginal von Experten-Concept-Maps profitieren (siehe 6.3.2 auf Seite 128). Das wurde bereits von Nesbit diskutiert (2006). Ergänzend zu O'Donnell hat sich herausgestellt, dass sich Experten-Concept-Maps nicht nur als Kommunikationshilfen für verbal schwächere Schüler eignen, sondern auch zu besseren Ergebnissen führen als vergleichbare Fachtexte (O'Donnell et al. 2002).

In Anlehnung an den Expertise-Reversal-Effekt²⁰ muss an dieser Stelle erwähnt werden, dass durch die Intervention keine messbare Verschlechterung in der Leistung bei leistungsstärkeren Schülern aufgetreten ist (Kalyuga et al. 2003).

Hieraus lässt sich folgende Empfehlung für die Schule ableiten: Im Fachunterricht sollten generell mehr Experten-Concept-Maps eingesetzt werden, als dies bisher geschieht. Für die gezielte individuelle Förderung von leistungsschwächeren Schülern sind Experten-Concept-Maps nur bedingt geeignet. Aus diesem Grund sollten Experten-Concept-Maps nicht als Pflicht-, sondern besser als Wahlangebot an Schüler gemacht werden. So können diese eine für sie persönlich passende Entscheidung in Bezug auf ihre Lernmaterialien treffen.

²⁰ Der Expertise-Reversal-Effekt bezieht sich auf die Umkehr der Wirksamkeit einer Intervention auf Lernende mit unterschiedlichem Vorwissen (Kalyuga 2007). Was für leistungsschwächere Schüler lernförderlich sein kann, kann für leistungsstärkere Schüler zu einer Hürde werden.

7.4 Zu den Genderaspekten

Jungen haben schlechtere Noten, aber keine schlechteren Ergebnisse

In den frühen 80ern hat sich die Geschlechterrate in Bezug auf die Schulnote umgekehrt (Becker und Müller 2011, Sievert und Kröhnert 2015). Schülerinnen erzielen seitdem bessere Noten als Schüler und das nicht nur in Deutschland (Prenzel 2007, OECD 2009, Quenzel und Hurrelmann 2010). Soziologen erklären diesen markanten Effekt zwischen den Geschlechtern durch die sprachliche Überlegenheit der Mädchen und mit nicht wünschenswertem Verhalten von Jungen im Unterricht. Während Mädchen im Unterricht tendenziell mehr Fleiß und Hingabe zeigen, tendieren Jungen eher zu Faulheit und Unordentlichkeit. Verstärkt werden diese geschlechtlichen Rollenmuster zudem durch Gruppendruck, was letztendlich zu schlechteren Noten bei den Jungen führt (Lupatsch und Hadjar 2011). Im Folgenden werden weitere Studien zu den Unterschieden zwischen den Geschlechtern aufgeführt, die Auswirkungen auf den schulischen Erfolg und Misserfolg haben:

- **Arbeitshaltung:** Mädchen beschreiben sich im Vergleich zu Jungen bei ihrer Arbeit in der Schule als gewissenhafter mit mehr Bereitschaft zu Anstrengungen (Maaz et al. 2011).
- **Anerkennung:** Im Gegensatz zu Jungen steht schulischer Erfolg bei Mädchen nicht im Widerspruch zu der Beliebtheit bei Mitschülern (Warrington et al. 2000).
- **Hausaufgaben:** Mädchen investieren mehr Zeit in die Erledigung der Hausaufgaben (OECD 2015).
- **Lesen:** Lesen korreliert mit dem Verstehen (Holsanova et al. 2009, S. 1224). So konnte bei den Pisa-Studien (2003, 2006, 2009 und 2012) festgestellt werden, dass Jungen zwar besser in Mathematik abschneiden, die Mädchen jedoch einen großen Vorteil beim Lesen haben (OECD 2015). Es hat sich aber auch gezeigt, dass besonders jene Schüler (unabhängig vom Geschlecht), die sich in ihrer Freizeit mit Lektüre beschäftigten, deutlich bessere Leistungen beim Lesen erzielten als Lesemuffel im gleichen Alter (Sievert und Kröhnert 2015). Aus diesem Grund geht die OECD davon aus, dass Jungen ihren Rückstand beim Lesen verbessern könnten, wenn sie in ihrer Freizeit öfter als bisher zu einem Buch greifen würden (OECD 2015).
- **Motivation:** Mädchen zeigen ein höheres Engagement bei schulischen Angelegenheiten als Jungen. Sie arbeiten infolgedessen härter für den schulischen Erfolg (Fischer et al. 2013).
- **Disziplin:** Neben der höheren Motivation zeigen Mädchen bei schulischen Dingen mehr Selbstdisziplin (Weis et al. 2013).

- Scheitern: Wenn sich bei Jungen kein Erfolg in der Schule einstellt oder sie an Aufgaben scheitern, tendieren sie dazu, die Schuld nicht bei sich selbst, sondern bei anderen oder den Rahmenbedingungen zu suchen. Mädchen empfinden Misserfolg tendenziell eher als persönliches Versagen (Ruble und Martin 1998).
- Selbstvertrauen: Geringes Selbstvertrauen in die eigenen Fähigkeiten kann den Erfolg in einem Schulfach auf negative Weise bestimmen (Sievert 2015).
- Biologische Unterschiede: Neben den bisher genannten Gründen werden immer wieder biologische Argumente aufgeführt. Im Prinzip drehen sich diese immer um die folgenden drei Themen: Gene, Hormone und Gehirn (Halpern 2012).
- Östrogene: Die bekanntesten weiblichen Sexualhormone können die Sprachfähigkeit steigern und räumliche Fähigkeiten verringern (Hampson 1990, Hausmann et al. 2000).
- Testosteron: Das Hormon Testosteron kann bei Personen zu einer verringerten Kontrolle der Emotionen führen. Die verminderte Kontrolle kann wiederum zu einer reduzierten Sensibilität der betroffenen Person gegenüber ihrer Umwelt führen (Yildirim und Derksen 2012). Metaanalysen haben darüber hinaus gezeigt, dass ein leichter Zusammenhang zwischen Testosteron und Aggression besteht (Book et al. 2001). Wie sich Testosteron bei Personen auf ihre Aggressivität auswirkt, hängt jedoch maßgeblich damit zusammen, wie liebevoll die Erziehung eines Kindes war (Yildirim und Derksen 2012). Neben den eher negativen Auswirkungen, kann Testosteron aber auch zu einer Wettbewerbsorientierung und Steigerung der Risikofreude führen (Carre und McCormick 2008).

Der relativ schnelle Wechsel der Geschlechterrate in Bezug auf die Schulnoten lässt sich allerdings nur bedingt durch biologische Unterschiede erklären (Calvin et al. 2010, Sievert und Kröhnert 2015, S. 28). Unabhängig von schulischen Leistungen und Noten zeigt eine Studie zum Vergleich der Ergebnisse zweier Fähigkeitstests (SAT und ACT) von 1,6 Millionen Siebtklässler aus den Jahrgängen 1981 bis 2010, dass Jungen seit knapp 30 Jahren (entgegen sich ändernder Gesellschaften und Rollenmuster) einen stabilen Vorsprung in Mathematik haben und Mädchen kontinuierlich besser bei der mündlichen und schriftlichen Ausdrucksfähigkeit abschneiden (Wai et al. 2010, S. 417). Im Folgenden werden Punkte aufgelistet, die über die Biologie hinausgehen und auch Auswirkungen auf den schulischen Erfolg und Misserfolg haben können:

- Erwartungen und Erfolg: Eine Studie hat gezeigt, dass Jungen in Tests tatsächlich schlechter abschneiden, wenn man ihnen vorher sagt, dass sie in diesem Test in der Regel immer schlechter abschneiden als Mädchen (Hartley und Sutton 2013).
- Diagnoseinstrumente: Unterschiede im Aufgabenformat (zum Beispiel Freiantwort und Multiple-Choice-Aufgaben) können Einfluss auf die empfundene Schwierig-

keit und den daraus resultierenden Erfolg bei Jungen und Mädchen haben. Mädchen empfinden Multiple-Choice-Aufgaben tendenziell schwieriger als Jungen (Close und Shiel 2009, Europäische Kommission – Exekutivagentur Bildung 2010).

- Erziehung: Kinder aus einem Elternhaus mit ausgewogen verteilten Arbeiten im Haushalt zeigen tendenziell auch ein modernes Rollenverhalten (Huerta et al. 2011). Eltern behandeln ihre Söhne und Töchter generell nicht ungleich, verstärken in manchen Bereichen jedoch geschlechtsspezifisches Verhalten (zum Beispiel Freizeitaktivitäten ihrer Kinder). Dies transportiert unterschiedliches Rollenverhalten mit entsprechende Erwartungen und festigt diese (Else-Quest et al. 2012).
- Lehrkräfte: Mädchen scheinen im Unterricht mit einem Bonus zu starten. Falschantworten werden von Lehrkräften bei Mädchen tendenziell weniger beanstandet, während Jungen für Falschantworten eher kritisiert werden (Schultheis 2012). Dies scheint den Jungen bewusst zu sein, denn sie fühlen sich ungerecht behandelt. Mädchen werden ihrer Meinung nach von Lehrkräften von vornherein besser eingeschätzt als Jungen (Koch-Priewe et al. 2009). Im Gegensatz dazu werden störende Jungen von Lehrkräften jedoch weniger stark gemäßregelt als Mädchen, die den Unterricht in vergleichbarem Maße stören (Francis 2002).
- Bedrohung durch Stereotype (stereotype threat): Der Stereotype threat bezeichnet die Befürchtung von Angehörigen einer stigmatisierten sozialen Gruppe, aufgrund von negativen Stereotypen über diese Gruppe anders beurteilt oder behandelt zu werden oder durch eigenes Verhalten in bestimmten Situationen ein Stereotyp zu bestätigen (Steele 1997, Steele et al. 2002). Dieses Gefühl der Bedrohung kann zu einer tatsächlich messbaren Leistungsminderung führen (Dorsch Lexikon der Psychologie 2015). Entsprechend ihrer Zugehörigkeit zu einer sozialen Gruppe (zum Beispiel männliche Schüler) kann es bei Jungen zu einem Gefühl der Bedrohung kommen. Hierbei kann das (negative) Stereotyp von Jungen durch das Verhalten des jeweiligen Schülers bestätigt werden oder zu einer selbsterfüllenden Prophezeiung führen „Jungen können das nicht, deshalb kann ich das auch nicht“ (Hartley und Sutton 2013, Hannover 2008, Ryan und Ryan 2005).
- Sozial konstruiertes Geschlecht (Gender): Nach Sievert (2015) schauen sich Mädchen und Jungen ihr geschlechtstypisches Verhalten in ihrem sozialen Umfeld ab und versuchen dabei den sozialen Erwartungen entsprechend normal zu sein (siehe Abb. 43).

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 43: Das Entwickeln von geschlechtstypischem Verhalten (Sievert und Kröhnert 2015, S. 31).

Obwohl die Ergebnisse nicht neu sind, so sind sie doch bemerkenswert. In der hier vorliegenden Studie haben die Jungen zwar weder signifikant bessere noch signifikant schlechtere Ergebnisse als die Mädchen erzielt, dennoch spiegelt sich ihre Leistung nicht in ihren Schulnoten wieder (siehe 6.4 ab Seite 129). Dieser Umstand ist unter anderem dadurch zu erklären, dass die (kognitive) Leistungsfähigkeit weit mehr ist als eine Schulnote (Helbig 2012).

Trotz alledem bilden Noten die Grundlage für die Vergabe von Studienplätzen, Jobs oder Praktika. Das ist nicht nur aus der Perspektive der Jungen unfair. Es widerspricht auch dem gesellschaftlichen Interesse alle Heranwachsenden bestmöglich zu bilden (Grundgesetz für die Bundesrepublik Deutschland, Spörhase 2012b). Nur so kann der Lebensstandard unserer Gesellschaft erhalten und enorme Folgekosten durch unzureichende Bildung vermieden werden (Piopiunik und Wößmann 2011, Woessmann und Piopiunik 2009). Dieser Umstand erfordert ein Überdenken von Noten bei der Beschreibung von Schülerleistungen.

Entsprechend den Befunden dieser Arbeit und den Befunden von Sievert (2015) wird deutlich, dass die fachdidaktische Forschung geschlechtsspezifische Unterschiede viel stärker als bisher berücksichtigen sollte. Weitere Unterschiede zwischen Jungen und Mädchen in Bezug auf die Hilfen werden unter 7.6.1 ab Seite 184 genannt und diskutiert, sowie Lösungsansätze formuliert.

7.5 Zu den Ergebnissen zum Faktor Zeit

7.5.1 Zeit als wirksamer Differenzierungsfaktor

Die Ergebnisse dieser Studie haben gezeigt, dass die Lernzeit ein hoch signifikanter Prädiktor für das Ergebnis von leistungsschwächeren Schülern ist (siehe 6.5.3 auf Seite 136). Im Allgemeinen besitzt die Lernzeit eines Schülers keinen Vorhersagewert auf das daraus resultierende Lernergebnis, da die Lernzeit an sich keinerlei Auskunft über die Qualität und die Intensität des Lernens gibt (Hesse 1994, Plant et al. 2005). In der Tat können Schüler zwar in der gleichen Zeit lernen, aber je nach Lernwirksamkeit kann das Ergebnis im Verhältnis zum Ausmaß der Zeit ganz anders ausfallen (Kamp et al. 2012). Aus diesem Grund plädiert van den Hurk bei Lernenden nach qualitativen Unterschieden zu suchen, um die abweichende Auswirkung von Zeit auf Individuen zu erklären (van den Hurk et al. 1998). In Bezug auf Carroll (Carroll 1963) nennt van den Hurk drei zusätzliche Variablen, die bei der Beschreibung von Schülerleistungen in Verbindung mit Zeit berücksichtigt werden müssen. Diese sind: die Zeit, die zum Lernen gegeben wird, die Zeit, die sich zum Lernen genommen wird und die Zeit, die zum Lernen benötigt wird (van den Hurk et al. 1998). Die vorliegende Studie untersuchte ausschließlich, wie viel Zeit sich Schüler für das Lernen genommen haben. Die Schüler bekamen in den Interviews so viel Zeit, wie sie wollten und hatten keine Möglichkeiten, anderen Aktivitäten nachzugehen. Im Gegensatz zu anderen Studien wurde in der vorliegenden das Leistungsniveau der Schüler berücksichtigt. Hierbei bestätigte sich, dass vor allem leistungsschwächere Schüler von mehr Zeit profitieren. Der Befund unterstützt Kalyugas Anmerkungen, dass vor allem Schüler mit geringen Vorkenntnissen und Fähigkeiten durch ein langsames Unterrichtstempo einen größeren Lernzuwachs erzielen (Kalyuga 2007).

Anders formuliert: Je mehr Zeit sich leistungsschwächere Schüler für die Erarbeitung nehmen, desto besser werden deren Ergebnisse. Im Umkehrschluss bedeute dies, dass gerade leistungsschwächere Schüler besonders stark unter unterrichtsbedingtem Zeitdruck leiden und zudem noch geringere Fortschritte als leistungsstärkere Schüler machen. Die hieraus resultierende Empfehlung für die Schule ist eine konsequente Differenzierung nach Zeit. Lehrkräfte sollten mehr Zeit für leistungsschwächere Schüler einplanen. Ferner sollten sie leistungsschwächere Schüler dazu ermutigen, sich ausreichend Zeit für die Erarbeitung zu nehmen, ohne sie dabei zu stigmatisieren (langsamer gleich dümmer). Um Leerlauf und Langeweile zu vermeiden, müssten für leistungsstärkere Schüler zusätzliche Aufgaben angeboten werden, damit alle Schüler einer Klasse die gleiche Chance auf lernwirksame Unterrichtszeit erhalten.

7.5.2 Leistungsschwächere Schüler zu fördern erfordert mehr Zeit von Lehrkräften

Interessanterweise lösten bei individueller Betreuung sowohl leistungsschwächere als auch leistungsstärkere Schüler die Aufgaben annähernd in der gleichen Zeit (siehe 6.5.2 ab Seite 133). Betrachtet man allerdings die Verteilung der Sprechzeiten während der Aufgaben, so zeigt sich, dass leistungsstärkere Schüler bei der Bearbeitung der Aufgaben mehr kommunizieren. Ferner musste der Versuchsleiter mit leistungsstärkeren Schülern weniger sprechen als mit leistungsschwächeren.

Als Empfehlung lässt sich für den Unterricht sagen, dass bei innerer Differenzierung über Leistung eine längere Begleitung von leistungsschwächeren Schülern durch die Lehrkraft in Arbeitsphasen eingeplant werden sollte, als im Vergleich zu leistungsstärkeren Schülern. Um dabei eine Vernachlässigung von leistungsstärkeren Schülern zu vermeiden, sollten deren Arbeitsprozesse und Ergebnisse trotzdem kontrolliert werden. Da Lehrkräfte im Unterricht zeitlich mehr an leistungsschwächere Schüler gebunden sind, sollte die Kontrolle der Ergebnisse durch Beispiel- oder Musterlösungen erfolgen. Diese würden die Selbstkontrolle durch leistungsstärkere Schüler ermöglichen und Lehrer entlasten.

7.6 Zu dem Einsatz der Hilfen bei den Lernaufgaben

7.6.1 Gestufte Hilfen bilden Verstehenschwierigkeiten ab und machen Unterschiede zwischen Schülern deutlich

Durch den Einsatz der Hilfen konnten bei den Schülern fachliche Schwierigkeiten bei den Aufgaben 1 bis 4 identifiziert werden (siehe Tab. 35 und Seite 147 ff.), die ein Ranking der Schwierigkeiten ermöglichten (siehe 6.7 Seite 146). Während der Interviews konnten darüber hinaus keine weiteren Verstehenschwierigkeiten gefunden werden, das heißt, die mit den Hilfen rekonstruierten kognitiven Operationen erwiesen sich als geeignet, mögliche Verstehenschwierigkeiten zu thematisieren.

Ausgehend von dem Ranking können die Schwierigkeiten der Schüler Ansatzpunkte für Unterrichtsplanung sein, denn gerade hier werden auch im Unterricht die meisten Verstehensprobleme auftreten. So zeigte sich zum Beispiel, dass Schüler mehr nicht-fachliche Hilfen bei der Aufgabenstellung benötigten, je weiter die Aufgaben von den Alltagserfahrungen der Schüler entfernt sind (siehe Tab. 35 auf Seite 147). Umgekehrt kann gesagt werden, dass je näher eine Aufgabe an den Alltagserfahrungen der Schüler liegt, desto weniger zusätzliche Hilfen werden zum Aufgabenverständnis benötigt.

Die Unterrichtsplanung kann demnach sowohl auf fachlicher als auch auf nichtfachlicher Ebene stattfinden. Hierbei kann die Lehrkraft entsprechend der Anforderungen (Unterschiede) ihrer Schüler Schwerpunkte setzen und im Vorfeld üben. So könnte bei lernschwächeren Schülern zum Beispiel gezielt das Interpretieren von Abbildungen geübt werden, bevor die Auseinandersetzung mit den Materialien dieser Studie stattfindet.

Weiter Diskussionspunkte zu den Ergebnissen in Bezug auf Gender

In Bezug auf die Unterschiede zwischen Schülergruppen hat sich bei dem Vergleich der Hilfen ferner gezeigt, dass Mädchen bei den Aufgaben 3 und 4 deutlich mehr Hilfe als Jungen benötigen (siehe Abb. 35 auf Seite 152). Das Gleiche zeigt sich bei dem Vergleich der Qualität der Schülerantworten (siehe Abb. 40 auf Seite 165).

Hierbei ist besonders bemerkenswert, dass die für Mädchen schwierigen Fachinhalte von Aufgabe 3 und 4 (Anwendung und Transfer) schon in den reproduktiven Aufgaben 1 und 2 thematisiert und dort bereits bewältigt wurden. Diese Inhalte führten bei den Aufgaben 1 und 2 zu keinen signifikanten Unterschieden und sind, bis auf den Lichteinfall ins Auge, reine Wiederholung von bereits Gelerntem. Was macht also den Unterschied?

- Zum einen könnte das bei Aufgaben 3 und 4 der erwähnte Lichteinfall ins Auge sein. Die Lösung der beiden Aufgaben gelingt besser, wenn man sich den Lichteinfall ins Auge plastisch vorstellen oder diesen am Augenmodell nachvollziehen kann. Da dies den Mädchen weniger gut gelungen ist als den Jungen, könnte das bedeuten, dass den Mädchen räumliche und plastische Vorstellungen schwerer fallen als den Jungen. Dies wird auch durch andere Studien belegt (Hausmann et al. 2000, Ma 2007, S. 79, Tzuriel und Egozi 2010).
- Zum anderen könnte dieser Unterschied aber auch durch die Art der Aufgaben entstehen. Die Aufgaben 3 und 4 gehen über die Auswahl und Reproduktion von Wissen hinaus und erfordern dessen Anwendung in einem unbekanntem, naturwissenschaftlichen Kontext in komplexerer Form als Aufgabe 1 und 2 (siehe die Einordnung des Schwierigkeitsgrades der Aufgaben 1 bis 4 entsprechend der KMK und dem Kompetenzstrukturmodell nach Kauertz auf Seite 113). Dies deckt sich mit der Einschätzung der Aufgaben durch die Schüler (siehe Kapitel 6.8.1 auf Seite 154 ff. und Tab. 37). Mädchen empfanden die ersten beiden Aufgaben tendenziell leichter als die letzten beiden Aufgaben. Bei den Jungen kehrt sich die Einschätzung um. Jungen empfanden die ersten beiden Aufgaben tendenziell schwieriger und die letzten beiden leichter. Das kann an den unterschiedlichen Interessen der Geschlechter in Bezug auf den Kontext der jeweiligen Aufgabe liegen

(Hoffmann 1997, Hammann 2006, S. 94, Scholz 2012, S. 9). So dreht es sich kontextuell bei Aufgabe 1 um den eigenen Körper und bei Aufgabe 2 um Tiere (Katzen). Bei Aufgabe 3 wird nach technischen Lösungen gefragt und bei Aufgabe 4 dreht es sich um das Erlebnis eines Jungen. Neben den unterschiedlichen Interessen könnte auch das unterschiedliche Selbstvertrauen der Geschlechter in die eigenen Fähigkeiten in den Naturwissenschaften eine Rolle spielen (Sievert und Kröhnert 2015, S. 22). Je mehr sich Schüler in einem Fach zutrauen, desto wahrscheinlicher bringen sie dort auch bessere Leistungen.

In Anlehnung an die geschlechtsspezifischen Ergebnisse dieser Arbeit und die in Kapitel 7.4 (S. 179) genannten Unterschiede zwischen den Geschlechtern werden im Folgenden Lösungsansätze für den Unterricht aufgeführt und stichwortartig in Abb. 44 auf Seite 187 zusammengefasst:

- **Multikausalität:** Zu Beginn muss gesagt werden, dass es bei der Geschlechterfrage keine einfache Antwort geben kann, da sich die Geschlechtsunterschiede (siehe 0) nicht auf einen oder wenige Einzelfaktoren reduzieren lassen (Halpern 2007). Daher muss jede Maßnahme immer unter Berücksichtigung der entsprechenden Lerngruppe getroffen werden.
- **Interessen bedienen:** Es ist möglich mehr Schülerinnen für die naturwissenschaftlichen Fächer zu begeistern und damit zu besseren Leistungen zu führen, wenn der Unterricht stärker die Interessen von Mädchen berücksichtigt (Hoffmann 1997).
- **Koedukation (reflexive):** Gemeinsamer Schulunterricht soll es Mädchen und Jungen unabhängig von geltenden Geschlechterstereotypen ermöglichen, ihre persönlichen Stärken zu entwickeln. Für den Unterricht erfordert das die Wahrnehmung und Förderung von individuellen Fähigkeiten jenseits von zugewiesenen Geschlechterrollen (Dorsch et al. 2014). Koedukation wird erleichtert, da keines der beiden Geschlechter eine bestimmte Unterrichtsform bevorzugt oder speziell darunter leidet (Halpern 2012). Darüber hinaus können geschlechtsbedingte Schwächen durch gezielte Maßnahmen überwunden werden, was an der Verbesserung der räumlichen Wahrnehmung bei Mädchen gezeigt werden konnte (Tzuril und Egozi 2010).
- **Lehrkräfte:** In Bezug auf Leistungsunterschiede von Lehrkräften in Abhängigkeit zu deren Geschlecht kann keine eindeutige Aussage gemacht werden. Fest steht aber, dass Lehrkräfte einen sehr großen Einfluss auf das Lernen aller Schüler haben und dass die größten messbaren Leistungsdifferenzen im Unterricht durch unterschiedlich gute Lehrkräfte entstehen (Hattie 2009). Dennoch ist es ratsam, dass jede Lehrkraft immer wieder ihr Handeln reflektiert und besonders darauf achtet Mädchen und Jungen nicht bevor- oder benachteiligt werden (Francis 2002).

- Notengebung: Leistungsdifferenzierung durch Lehrkräfte beachtet Begabung und Fähigkeiten, soziale Merkmale sowie motivationale Aspekte und Tugenden (Anders et al. 2010, S. 327). Zum Wohle der Jungen sollte über die bisherige Leistungsdifferenzierung hinaus besonders bei der Vergabe von Zeugnis- und Abschlussnoten weniger frei entschieden, sondern mehr auf Standardisierung und Formalisierung von Prüfungen geachtet werden (Diefenbach 2012).

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 44: Lösungsansätze zum Umgang mit Geschlechterunterschieden im naturwissenschaftlichen Unterricht (nach Sievert und Kröhnert 2015, S. 39).

7.6.2 Ein Viertel aller Schüler benötigt nicht-fachliche Unterstützung

Der Umgang der Schüler mit den Hilfen offenbarte, dass knapp ein Viertel aller Schüler an nicht-fachlichen Anforderungen einer Aufgabe scheiterten (siehe 6.6.2 ab Seite 145). Dies stimmt mit aktuellen Ergebnissen der PISA Studie 2012 überein (OECD 2014). Hier hat sich beim internationalen Vergleich der OECD-Länder gezeigt, dass etwa ein Fünftel aller Schüler maximal nur einfache Probleme lösen können, die sich ausschließlich auf vertraute Situationen beziehen. Aufgaben mit unvertrautem Kontext oder gar Transfer waren für diesen Teil der Schüler der PISA Studie ohne zusätzliche Hilfe nicht lösbar.

Das Scheitern der Schüler an nicht-fachlichen Problemen sollte zur Konsequenz haben, dass im Unterricht weiterhin fachliche Unterstützung gegeben werden muss. Diese sollte aber durch bedarfsgerechte und nicht-fachliche Unterstützung erweitert werden. Entsprechende Erweiterungen und Unterstützungen sollten sich vor allem auf die Förderung der Sprach- und Lesekompetenz von leistungsschwächeren Schülern sowie Schülern mit Migrationshintergrund beziehen (Stanat et al. 2010, Köller 2012, PISA 2012 Ergebnisse 2013).

Zur Steigerung der Qualität von Lernaufgaben im Unterricht und gegen das Scheitern an nicht-fachlichen Problemen genügt es nicht, Schüler etwas erarbeiten zu lassen, ohne sie darüber aufzuklären, was sie tun, wie sie es bewerkstelligen können und wozu es geschehen soll (Rademann und Blume 2001, Thonhauser 2008). Das gilt sowohl für die Lern- als auch für die Leistungsphasen im Unterricht. Es ist unbestritten, dass es zu keinem Lernzuwachs kommt, wenn Schüler nicht verstehen, was sie erarbeiten sollen. Zudem handelt es sich nicht um eine objektive Leistungsmessung, wenn Schüler nicht verstehen, was von ihnen in einer Testsituation erwartet wird (Müller und Helmke 2008).

Beim Entwickeln von nicht-fachlichen Hilfen geht es hierbei als Lehrkraft also vor allem darum, die eigenen Aufgaben, das eingesetzte Material oder die Methoden aus Sicht der Schüler neu zu überdenken, nicht-fachliche Stolpersteine und Verstehensschwierigkeiten vorauszudenken und diese durch entsprechende Hilfen zu entschärfen. Ein solches Vorausdenken fordert von den Lehrkräften ein fundiertes fachliches Inhaltswissen (CK), pädagogische Kenntnisse (PK) über lern- und medienpsychologische Grundlagen sowie fachdidaktisches Wissen (pedagogical content knowledge, PCK) (Shulman 1986, Strunz-Maireder 2009).

Derartige Überlegungen können von Lehrkräften nicht immer zu 100% im Voraus geschehen, denn oft überrascht die Unterrichtsrealität mit neuen Facetten, die im Vorfeld nicht bedacht wurden. Nichtsdestotrotz sollten die Erfahrungen nach dem Unterricht

reflektiert und zumindest als Notiz in die eigenen Unterrichtsmaterialien einfließen, damit die gleichen Stolpersteine nicht erneut auftreten. Die Reflektion und Überarbeitung²¹ des Unterrichts zeigt auch, wie wichtig die Nachbereitung ist und wie wenig Zeit dafür zur Verfügung steht (Nido et al. 2008).

7.7 Der Schwierigkeitsgrad von Aufgaben ist planbar

Die Staffelung entsprechend des Schwierigkeitsgrades bei der Konstruktion der Aufgaben ist gelungen (siehe Tab. 18 auf Seite 113) und wurde von den Schülern bestätigt (siehe Tab. 36 auf Seite 155).

Das zeigt, dass unter Berücksichtigung theoretischer Aufgabenmerkmale, wie zum Beispiel des Kompetenzstufenmodells für die naturwissenschaftlichen Fächer (Mayer et al. 2013), Aufgaben gezielt für unterschiedliche Schwierigkeitsgrade konstruiert werden können. Dies ermöglicht eine fundierte Planung von gestaffelten Aufgabenreihen zu Themen der Biologie über unterschiedliche Schwierigkeitsgrade hinweg. Die gezielte Konstruktion könnte nicht nur den Mangel an Transferaufgaben verringern (Hamman 2006, Germ und Harms 2010), sondern würde auch die Aufgabenkultur im Biologieunterricht durch mehr Vielfalt und Bandbreite an Aufgaben erweitern (Thonhauser 2008, Kühn 2011).

Ferner hat die deckungsgleiche Einschätzung der Aufgabenschwierigkeit durch die Schüler gezeigt, dass sich Schüler der Schwierigkeit von Aufgaben bewusst sind. Die Fähigkeit, Aufgaben einzuschätzen, ermöglicht bei innerer Differenzierung im Biologieunterricht eine Auswahl passender Lernaufgaben durch die Schüler selbst (selbstdifferenzierender Ansatz). Die Selbstauswahl entkräftet dabei das kritische Zuteilen von Aufgaben durch die Lehrkraft, das mit Fehleinschätzungen und Stigmatisierungen einhergehen kann (siehe Tab. 4 auf Seite 30 und Scholz 2008).

7.8 Der Schwierigkeitsgrad einer Aufgabe fördert Interesse

Ein bemerkenswertes Ergebnis ist nicht nur die deckungsgleiche Einschätzung des Schwierigkeitsgrades der Aufgaben sondern auch, dass die meisten Schüler die schwierigste Aufgabe am interessantesten fanden (siehe Tab. 36 auf Seite 155). Ferner zeigten die Argumente der Schüler zur Unterstützung ihrer Einschätzungen der Aufgaben Übereinstimmungen. Die Kongruenz der Argumente bezieht sich vor allem auf das, was eine Aufgabe nach Schülermeinung schwierig macht und auf das, was eine

²¹ Als Kurzhinweis zur Überarbeitung der eigenen Materialien führt Prediger folgende Fragen aus Sicht der Schüler auf (Prediger 2008, S. 6): „Was kann ich in dieser Einheit lernen? Was muss ich dazu erstellen, was muss ich zeigen? Was muss ich tun, um dorthin zu gelangen? Wo finde ich das, was ich zum Arbeiten brauche? Was kann ich schon, was noch nicht?“

Aufgabe interessant macht (siehe 6.8.3 ab Seite 157 und 6.8.4 ab Seite 158). Die Argumente der Schüler bekräftigen bestehende Desiderata für kumulatives Lernen im Biologieunterricht. Die Desiderata besagen, dass bedeutungsvolles Lernen einen gewissen Grad an Komplexität benötigt, also nicht zu einfach sein soll, dass es an lebensweltliche Erfahrungen anknüpft, auf die naturwissenschaftliche Erklärungen folgen, dass Zusammenhänge klar werden und eine Anwendung von Gelerntem stattfindet (Kattmann 2003).

Als Plädoyer für die Schule lässt sich hieraus formulieren, dass Aufgaben im Biologieunterricht nicht nur im Bereich der Minimalstandards liegen, sondern auch darüber hinausgehen sollten. Dies lässt sich damit begründen, dass trotz der individuell unterschiedlich empfundenen Schwierigkeit der Aufgaben annähernd alle Schüler die schwierigeren Aufgaben auch interessanter fanden, weil sie gerade bei diesen Aufgaben den größten persönlichen Lernzuwachs feststellen konnten (Leutner et al. 2008).

7.9 Schwierigkeiten der Schüler bei den Abbildungen

Medien zu lesen muss geübt werden

Die Ergebnisse dieser Studie zeigen, dass sowohl leistungsschwächer Schüler (siehe Abb. 33 auf Seite 150), als auch Schüler, die bereits mit dem Fachtext und/oder der Experten-Concept-Map Probleme hatten (siehe Abb. 36 auf Seite 153), sich ebenfalls bei der Benennung der Inhalte in den Abbildungen schwer taten. Das führte dazu, dass diese beiden Schülergruppen annähernd automatisch bei Aufgabe 1 scheiterten. Da Aufgabe 1 die Basis für das Lösen der folgenden Aufgaben bildet, wirkte sich ein Scheitern an dieser Aufgabe auch negativ auf das Lösen der nachfolgenden Aufgaben aus.

Wie bei anderen Repräsentationen auch, hat sich bei der Benennung der Abbildungen in den Interviews herausgestellt, dass je abstrakter der Inhalt war, umso schwerer fiel es den Schülern, die Abbildung und deren Inhalt zu beschreiben (siehe 6.8.6 ab Seite 167). Die Abstraktion der Abbildungen (siehe 9.1.4 auf Seite 224) entstand unter anderem durch die unterschiedlichen Vergrößerungsgrade der Abbildungen. Es wurden so verschiedene biologische Ebenen dargestellt, was Schülern bekanntermaßen schwer fällt (Bahar et al. 1999). Dies führte zu folgenden Problemen:

- 1) Die Lernenden müssen zuerst das verwendete Symbolsystem verstehen, bevor sie die damit transportierten Inhalte begreifen (Krapp 2006, S. 435). Gelingt ihnen das nicht, fehlt ihnen ein Teil der Informationen. Mangelndes Bildverstehen verhält sich in diesem Falle so, als ob die Schüler beim Lesen eines Textes einen Teil übersprungen hätten.

- 2) Sobald im Unterricht mehr als ein Medium zum Wissenstransport verwendet wird und das ist in der Biologie der Normalfall (Schwanewedel und Nitz 2013, S. 11) steigert sich die Komplexität und damit der Schwierigkeitsgrad. Sofern die verwendeten Medien dem Redundanzprinzip von Mayer (Mayer 2001) entsprechen, transportieren Repräsentationen nicht die gleichen Inhalte, sondern ergänzen sich gegenseitig. Daher müssen Lernende nicht nur den Inhalt des jeweiligen Mediums verstehen, sondern auch deren Relation untereinander erkennen (Brünken et al. 2005, S. 62). Falls Schülern jene Verknüpfung nicht gelingt, fehlt ihnen nicht nur die Information des jeweiligen Mediums, sondern auch deren Zusammenhang (Lachmayer et al. 2007, Klöpfel et al. 2013).

Für die Umsetzung im Unterricht lassen sich folgende Empfehlungen formulieren:

Wenn

- mehr als ein Medium im Unterricht eingesetzt wird, sollte deren Zusammenhang klar sein. Das lässt sich zum Beispiel in der Arbeitsanweisung verankern oder durch die Verwendung von gleichen Farben, verständlichen Legenden oder aussagekräftigen Beschriftungen verdeutlichen.
- ein Medium eingesetzt wird, das den Schülern bisher wenig bekannt ist, sollte im Vorfeld die Herangehensweise und der Umgang mit dem entsprechenden Medium geklärt und nicht nur am Ende dessen Inhalt besprochen werden. Im Fall der Erarbeitung mit Experten-Concept-Maps wären das zum Beispiel der nicht existierende Start, die nicht vorhandene Leserichtung, sowie Pfeile mit Sackgassen und dem daraus resultierenden Rückwärtslesen.

Je nach Bekanntheitsgrad der Medien müssen Lernende sich intensiv mit den jeweiligen Repräsentationen auseinandersetzen, um eine angemessene Verarbeitungstiefe zu erreichen (Krapp 2006, S. 438 f.). In Abhängigkeit des Niveaus und dem Vorwissen der Lernenden sinkt oder erhöht sich dabei der Bedarf der Schüler an mehr Bearbeitungszeit und mehr Führung durch das Material (Kalyuga 2007).

7.10 Limitationen der Studie

Landesbestimmungen und Testverfahren

Zur Unterscheidung zwischen leistungsstärkeren und leistungsschwächeren Schülern haben wir uns für die Biologienote als Kriterium entschieden. Dies geschah aus zwei Gründen. Zum einen sollte die Biologienote Auskunft über das Wissen und Können der Schüler im Bereich biologischer Themen geben. Zum anderen war es uns aus Gründen der Gesetzgebung des Landes Baden-Württemberg nicht gestattet, Intelli-

genz- oder Fähigkeitstest im Rahmen dieser Studie durchzuführen. Dies ist in Baden-Württemberg ausschließlich Psychologen erlaubt, um auf diese Weise Missbrauch und Fehlinterpretationen zu vermeiden.

Anzahl der Schüler

Eine weitere Limitation dieser explorativen Studie war die Anzahl der Teilnehmer. Die Anzahl von 112 Personen war zwar ausreichend für die Überprüfung auf Haupt- und Interaktionseffekte bei der Hauptgruppe von Schülern. Sobald jedoch die gleichen Berechnungen für Untergruppen durchgeführt werden sollten, wurde die Anzahl der Testpersonen in Bezug auf die jeweiligen Voraussetzungen der Tests kritisch.

Die Reduzierung von 291 angefragten Schülern auf 112 Teilnehmer lag an dem krankheitsbedingten Ausfall einer teilnehmenden Kollegin. Die eingesetzte Vertretungslehrkraft verweigerte die Teilnahme der Klassen. In Anbetracht des Fortschreitens der Studie und der benötigten Vorlaufzeit einer erneuten Rekrutierung und Durchführung der Interviews, konnte dieser Wegfall nicht kompensiert werden.

Unterschiede zwischen Reproduktions- und Transferaufgaben

In dieser Studie wurden nach einer kurzen Einführung vier unterschiedliche Lernaufgaben mit Schülern durchgeführt. Bei der Durchführung handelte es sich um eine Reihung von Aufgaben zum Aufbau der Netzhaut bei Menschen mit steigendem Schwierigkeitsgrad²². Hierbei hatten die Schüler weder Zeit zum Wiederholen oder Festigen der neuen Fachinhalte, noch eine Verschnaufpause zwischen den Aufgaben. Darüber hinaus setzten sich die Materialien überwiegend aus Papier und Stift zusammen. Daher ist es kaum verwunderlich, dass die Reihung der Aufgaben in so kurzer Zeit kaum messbare Auswirkungen bei den Transferaufgaben erbrachte. Lernen braucht Zeit und Übung (Klauer 2010, S. 17, Mandl et al. 2004, S. 13).

7.11 Abschließende Bemerkungen

Ein aus dieser Forschungsarbeit resultierendes Forschungsdesiderat zur Förderung von Differenzierung in der Schule ist die partizipative Entwicklung schultauglicher und diagnostischer Instrumente in Zusammenarbeit mit Lehrern. Diese Zusammenarbeit hat folgende Vorteile: Erstens würde die Zusammenarbeit das Bewusstsein von Lehrkräften gegenüber standardisierten Tests schärfen und zweitens die Chance auf die Umsetzung von Differenzierungsmaßnahmen im Schulalltag erhöhen.

²² Der Strahlengang des Lichtes durch den optischen Augenapparat wurde im Interview aus Zeitgründen nicht explizit thematisiert.

Differenzierung ist die Anpassung des Unterrichts an die Bedürfnisse der Schüler. Um den Unterricht diesen Bedürfnissen anzupassen, müssen diese der Lehrkraft klar sein. Aus diesem Grund sollten Diagnosen vielmehr auf Testergebnissen basieren und weniger auf Gefühl beruhen (siehe subjektive Fehlerquellen in Tab. 4 auf Seite 30). Aus rechtlichen Gründen wurde in dieser Studie die Schulnote zur Einteilung der Schüler eingesetzt. Die Schulnote ist allerdings kein geeignetes Instrument zur Diagnose im Unterricht, denn Noten beschreiben ein Symptom von Leistungsschwäche und nicht deren Ursache.

Gerade in Gesamtschulen ist der produktive Umgang mit Vielfalt besonders wichtig, da hier die Heterogenität am höchsten ist. Trotz der Möglichkeiten von innerer Differenzierung, haben auch diese ihre Grenzen. So sollten mit zunehmender Heterogenität zusätzliche Förderkurse für bestimmte Schüler in Betracht gezogen werden. Im Idealfall würde das eine Zusammenarbeit mit Sonderpädagogen bedeuten.

Experten-Concept-Maps zeigten positive Auswirkungen auf das Lernen von Schülern. Dies spricht für einen häufigeren Einsatz von (Experten-)Concept-Maps im Unterricht. Um deren Einsatz zu erleichtern, sollte ein Handbuch oder Online-Manual für die Erstellung, Anpassung und Nutzung von Experten-Concept-Maps in der Schule entworfen werden, so dass mehr Schüler von dieser Methode profitieren können. In einem späteren Schritt sollte dann deren Anwendung und Effekte in der Schulpraxis erneut getestet und bewertet werden.

8 QUELLENVERZEICHNIS

- Arbeitsgemeinschaft Externe Evaluation von Schulen (ARGEV). Online verfügbar unter <http://www.argev.ch/index.php>, zuletzt geprüft am 17.01.2014.
- Staatsinstitut für Schulqualität und Bildungsforschung (ISB). Online verfügbar unter <http://www.isb.bayern.de/>, zuletzt geprüft am 17.01.2014.
- (vom 11.11.1953): Verfassung des Landes Baden-Württemberg, S. 73–109. Online verfügbar unter http://www.landtag-bw.de/files/live/sites/LTBW/files/dokumente/rechtliche_grundlagen/Landesverfassung.pdf, zuletzt geprüft am 25.06.2013.
- SchG, vom 01.08.1983 (01.08.1983): Schulgesetz für Baden-Württemberg in der Fassung vom 1. August 1983, zuletzt geprüft am 13.03.2012.
- Naturwissenschaftliches Arbeiten 3 (2005). 1. Aufl., 1. Dr. Berlin: Cornelsen.
- PISA 2012 Ergebnisse. Was Schülerinnen und Schüler wissen und können; Programme for International Student Assessment (2013). Bielefeld: Bertelsmann.
- Anders, Yvonne; McElvany, Nele; Baumert, Jürgen (2010): Wie differenziert urteilen Lehrkräfte? In: Kai Maaz und Yvonne Anders (Hg.): Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten, Bd. 34. Berlin: Bundesministerium für Bildung und Forschung (BMBF) (Bildungsforschung, 34), S. 313–330.
- Astleitner, Hermann (2008): Die lernrelevante Ordnung von Aufgaben nach der Aufgabenschwierigkeit. In: Josef Thonhauser (Hg.): Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik. Münster: Waxmann, S. 65–80.
- Aufschnaiter, Claudia von (2003): Leitfaden zu den publizierten Ergebnissen der Forschungsarbeiten. Videobasierte Analysen von Lern- und Lehrprozessen in physikalischen Kontexten. Universität Hannover - Fachbereich Erziehungswissenschaften. Hannover. Online verfügbar unter http://www.cvauf.de/material/Habil_CvAuf.pdf, zuletzt geprüft am 14.06.2013.
- Aufschnaiter, Stefan von (2000): Kompliziert oder schwierig - kognitive Entwicklung beim Lösen von Aufgaben. In: Horst Bayrhuber (Hg.): Lehren und Lernen im Biologieunterricht. Innsbruck: Studien-Verl. (Forschungen zur Fachdidaktik, Bd. 1).
- Bahar, M.; Johnstone, A. H.; Hansell, M. H. (1999): Revisiting learning difficulties in biology. In: Journal of Biological Education (Society of Biology).

- Bayerl, Petra Saskia; Paul, Karsten Ingmar (2011): What Determines Inter-Coder Agreement in Manual Annotations? A Meta-Analytic Investigation. In: Computational Linguistics 37 (4), S. 699–725.
- Bayrhuber, Horst (Hg.) (2000): Lehren und Lernen im Biologieunterricht. Innsbruck: Studien-Verl. (Forschungen zur Fachdidaktik, Bd. 1).
- Bayrhuber, Horst; Bögeholz, Susanne; Elster, Doris; Hammann, Marcus; Hössle, Corinna; Lücken, Markus et al. (2007): Biologie im Kontext. Ein Programm zur Kompetenzförderung durch Kontextorientierung im Biologieunterricht und zur Unterstützung von Lehrerprofessionalisierung. In: Der mathematische und naturwissenschaftliche Unterricht : MNU; Organ des Deutschen Vereins zur Förderung des Mathematischen und Naturwissenschaftlichen Unterrichts e.V. 60 (5), S. 282–286.
- Becker, Hans-Jürgen (2009): Differenzierung - was ist gemeint? Ein Konstrukt im Spiegel chemiedidaktischer Zeitschriftenpublikationen. In: Naturwissenschaften im Unterricht. Chemie 20 (111/112), S. 5–7.
- Becker, Rolf; Müller, Walter (2011): Bildungsungleichheiten nach Geschlecht und Herkunft im Wandel. In: Andreas Hadjar (Hg.): Geschlechtsspezifische Bildungsungleichheiten: VS Verlag für Sozialwissenschaften, S. 55-75.
- Blayney, Paul; Kalyuga, Slava; Sweller, John (2015): The impact of complexity on the expertise reversal effect: experimental evidence from testing accounting students. In: Educational Psychology, S. 1–18.
- Blömeke, Sigrid; Risse, Jana; Müller, Christian; Eichler, Dana; Schulz, Wolfgang (2006): Analyse der Qualität von Aufgaben aus didaktischer und fachlicher Sicht. Ein allgemeines Modell und seine exemplarische Umsetzung im Unterrichtsfach Mathematik. In: Unterrichtswissenschaft Weinheim: Zeitschrift für Lernforschung 34 (4), S. 330–357.
- Bloom, Benjamin Samuel; Engelhart, Max D (1976): Taxonomie von Lernzielen im kognitiven Bereich. 5. Aufl., (17. - 21. Tsd.). Weinheim: Beltz (Beltz-Studienbuch, 35).
- Blume, Bernd; Rademann, Eva (2000): Ergebnisse einer Mängelanalyse der SINUS-Arbeitsgruppen des Sets 2 in Schleswig-Holstein, zuletzt geprüft am 07.07.2009.
- Bohl, Thorsten (2012): Umgang mit Heterogenität im Unterricht: Forschungsstand, Problembereiche, Perspektiven. Die Vielfalt des Lernens entdecken - Individuelle Förderung an beruflichen Schulen. Eberhard Karls Universität Tübingen. Leinfelden-Echterdingen, 2012.
- Bohl, Thorsten; Bönsch, Manfred; Trautmann, Matthias; Wischer, Beate (2012): Binnendifferenzierung. Ein altes Thema in der aktuellen Diskussion. Zur

- Einleitung. In: Beate Wischer (Hg.): Binnendifferenzierung. Teil 1: Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht. Unter Mitarbeit von Thorsten Bohl, Manfred Bönsch und Matthias Trautmann. neue Ausg. Immenhausen Hess: Prolog-Verlag (17), S. 5–7.
- Bönsch, Manfred (2012): Strategien zur Lernprozessoptimierung - Innere Differenzierung. In: Beate Wischer (Hg.): Binnendifferenzierung. Teil 1: Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht. Unter Mitarbeit von Thorsten Bohl, Manfred Bönsch und Matthias Trautmann. neue Ausg. Immenhausen Hess: Prolog-Verlag (17), S. 9–22.
- Book, Angela S.; Starzyk, Katherine B.; Quinsey, Vernon L. (2001): The relationship between testosterone and aggression: a meta-analysis. In: *Aggression and Violent Behavior* 6 (6), S. 579–599.
- Bortz, Jürgen; Döring, Nicola (2006): *Forschungsmethoden und Evaluation. Für Human- und Sozialwissenschaftler*. 4. Aufl. Heidelberg: Springer (Springer-Lehrbuch).
- Bräu, Karin (2007): Die Betreuung der Schüler im individualisierenden Unterricht der Sekundarstufe. Strategien und Handlungsmuster der Lehrenden. In: Kerstin Rabenstein und u.a. (Hg.): *Kooperatives und selbstständiges Arbeiten von Schülern. Zur Qualitätsentwicklung von Unterricht*.
- Bresler, Siegfried (2007): Mit Kompetenzrastern Unterricht planen und bewerten. In: Rita Wodzinski, Christoph Wodzinski und Ralph Hepp (Hg.): *Differenzierung. Naturwissenschaften im Unterricht Physik* 18 (99/100): Erhard Friedrich Verlag, S. 32–33.
- Bruder, Regina (2000): Mit Aufgaben arbeiten: Ein ganzheitliches Konzept für eine andere Aufgabenkultur. In: *Mathematik lehren: die Zeitschr. für d. Unterricht in allen Schulstufen* (101), S. 12–17.
- Bruder, Regina (2003): Konstruieren - auswählen - begleiten. Über den Umgang mit Aufgaben. In: *Friedrich-Jahresheft: Aufgaben*, S. 12–15.
- Brüning, Ludger; Saum, Tobias (2007): Sachtexthe verstehen - durch grafisches Strukturieren. In: *Deutschunterricht* 60 (4), S. 30–32.
- Brünken, Roland; Seufert, Tina; Zander, Steffi (2005): Förderung der Kohärenzbildung beim Lernen mit multiplen Repräsentationen. In: *Zeitschrift für Pädagogische Psychologie* 19 (1), S. 61–75.
- Budde, Jürgen (2012): Die Rede von der Heterogenität in der Schulpädagogik. Diskursanalytische Perspektiven. In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*; Vol 13, No 2 (2012): *Qualitative Computing: Diverse Worlds and Research Practices* 13 (2), S. 1–31.

- Bundesministerium für Unterricht, Kunst und Kultur (Hg.): Schulqualität Allgemeinbildung. Online verfügbar unter <http://www.sqa.at/>, zuletzt geprüft am 17.01.2014.
- Bundestag (21.12.2008): Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen sowie zu dem Fakultativprotokoll vom 13. Dezember 2006 zum Übereinkommen der Vereinten Nationen über die Rechte von Menschen mit Behinderungen. Fundstelle: <http://www.un.org/Depts/german/uebereinkommen/ar61106-dbgbl.pdf> 2008. Online verfügbar unter <http://www.un.org/Depts/german/uebereinkommen/ar61106-dbgbl.pdf>, zuletzt geprüft am 24.06.2013.
- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (1997): Gutachten zur Vorbereitung des Programms "Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts". Bonn: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (Materialien zur Bildungsplanung und zur Forschungsförderung, 60).
- Bybee, Rodger W. (1997): *Achieving scientific literacy. From purposes to practices.* Portsmouth, NH: Heinemann.
- Calvin, Catherine M.; Fernandes, Cres; Smith, Pauline; Visscher, Peter M.; Deary, Ian J. (2010): Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. In: *Intelligence* 38 (4), S. 424–432.
- Campbell, Neil A.; Kratochwil, Anselm; Lazar, Thomas; Reece, Jane B. (2009): *Biologie*. 8., aktualisierte Aufl. [der engl. Orig.-Ausg., 3. Aufl. der dt. Übers.]. München [u.a.]: Pearson Studium (Pearson Studium - Biologie).
- Caroll, John B. (1963): A Model of School Learning. In: *Teachers College Record* 64 (8), S. 723–733.
- Carre, Justin M.; McCormick, Cheryl M. (2008): Aggressive behavior and change in salivary testosterone concentrations predict willingness to engage in a competitive task. In: *Hormones and Behavior* 54 (3), S. 403–409.
- Chandler, Paul; Sweller, John (1991): Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*. In: *Cognition and Instruction* 8 (4), S. 293–332, zuletzt geprüft am 01.10.2012.
- Chang, K. E.; Sung, Y. T.; Chen, S. F. (2001): Learning through computer-based concept mapping with scaffolding aid. In: *Journal of Computer Assisted Learning* 17 (1), S. 21–33.

- Chou, Pao-Nan (2013): Effect Of Instructor-Provided Concept Maps And Self-Directed Learning Ability On Students' Online Hypermedia Learning Performance. In: *Journal of College Teaching & Learning* 10 (4), S. 223–233.
- Close, Sean; Shiel, Gerry (2009): Gender and PISA mathematics: Irish results in context. In: *European educational research journal* 1 (1), S. 20–33.
- Cohen, J. (1992): A power primer. In: *Psychol Bull* 112 (1), S. 155–159.
- Cohen, Jacob (1960): A coefficient of agreement for nominal scales. In: *Educational and Psychological Measurement*, S. 37–46.
- Czeskleba, Anja; Schmiemann, Philipp (2012): Steuerung des Kompetenzerwerbs beim Lernen mit Beispielaufgaben durch fokussierte Lernimpulse. Projektskizze. In: *Erkenntnisweg Biologiedidaktik* (11), S. 101–114.
- Davies, Martin (2011): Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? In: *Higher Education* 62 (3), S. 279–301.
- Diefenbach, Heike (2012): Jungen und schulische Bildung. In: Bernd Ahrbeck und Michael Matzner (Hg.): *Handbuch Jungen-Pädagogik*. [aktualisiert und mit neuen Beiträgen]. 2. Aufl. Weinheim [u.a.]: Beltz (Pädagogik), S. 109–126.
- Dohnicht, Jörg (2004): Medien im Unterricht. In: Ulrich Abele, Gislinde Bovet und Volker Huwendiek (Hg.): *Leitfaden Schulpraxis. Pädagogik und Psychologie für den Lehrberuf*. 4. Aufl. Berlin: Cornelsen, S. 150–176.
- Dorsch Lexikon der Psychologie (2015): Stereotype threat - Dorsch Lexikon der Psychologie - Verlag Hans Huber. Online verfügbar unter <https://portal.hogrefe.com/dorsch/stereotype-threat/>, zuletzt geprüft am 19.07.2015.
- Dorsch, Friedrich; Wirtz, Markus Antonius; Strohmer, Janina (Hg.) (2014): *Dorsch - Lexikon der Psychologie*. 17. Aufl. Bern: Huber.
- Else-Quest, Nicole M.; Higgins, Ashley; Allison, Carlie; Morton, Lindsay C. (2012): Gender differences in self-conscious emotional experience: A meta-analysis. In: *Psychological Bulletin* 138 (5), S. 947–981.
- Eppler, Martin J. (2006): A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. In: *Inf Vis* 5 (3), S. 202–210, zuletzt geprüft am 30.10.2013.
- Europäische Kommission-Exekutivagentur Bildung, Audiovisuelles und Kultur (2010): *Geschlechterunterschiede bei Bildungsergebnissen. Derzeitige Situation*

- und aktuelle Maßnahmen in Europa. In: Gender differences in educational outcomes, S. 1–158, zuletzt geprüft am 23.05.2014.
- Fey, Anja; Gräsel, Cornelia; Puhl, Thomas; Parchmann, Ilka (2004): Implementation einer kontextorientierten Unterrichtskonzeption für den Chemieunterricht. In: Unterrichtswissenschaft 32 (3), S. 238–256.
- Fischer, Franziska; Schult, Johannes; Hell, Benedikt (2013): Sex differences in secondary school success: Why female students perform better. In: European Journal of Psychology of Education 28 (2), S. 529–543.
- Fischer, Hans E.; Klemm, Klaus; Leutner, Detlev; Sumfleth, Elke; Tiemann, Rüdiger; Wirth, Joachim (2003): Naturwissenschaftsdidaktische Lehr-Lernforschung: Defizite und Desiderata. In: Zeitschrift für Didaktik der Naturwissenschaften 9, S. 179–209.
- Forschergruppe-Universität Kassel, unter der Leitung von: Hänze Martin; Stäudel, Lutz; Wodzinski, Rita (2007): Schritt für Schritt zur Lösung. Differenzierung durch Aufgaben mit gestuften Lernhilfen. Unter Mitarbeit von Gudrun Franke-Braun, Florian Schmidt-Weigand und Simone Blum. In: Rita Wodzinski, Christoph Wodzinski und Ralph Hepp (Hg.): Differenzierung. Naturwissenschaften im Unterricht Physik 18 (99/100): Erhard Friedrich Verlag, S. 42–45.
- Francis, Becky (2002): Boys, Girls and Achievement. Addressing the Classroom Issues. Hoboken: Taylor and Francis.
- Freelon, Deen (2013): ReCal OIR: Ordinal, Interval, and Ratio Intercoder Reliability as a Web Service. In: International Journal of Internet Science 8 (1), S. 10–16. Online verfügbar unter http://www.ijis.net/ijis8_1/ijis8_1_freelon_pre.html, zuletzt geprüft am 04.05.2014.
- Friedrichs, Jürgen; Lüdtke, Hartmut (1971): Teilnehmende Beobachtung. Zur Grundlegung einer sozialwissenschaftlichen Methode empirischer Feldforschung. Weinheim u.a: Beltz (Veröffentlichungen : Reihe E, Untersuchungen, 11).
- Friege, Gunnar (2001): Wissen und Problemlösen. Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs. Berlin: Logos (Studien zum Physiklernen, 19).
- Fromm, Sabine (2010): Datenanalyse mit SPSS für Fortgeschrittene. 1. Aufl. Wiesbaden: VS, Verl. für Sozialwiss. (Lehrbuch).
- Fuchs, Mandy (2006): Vorgehensweisen mathematisch potentiell begabter Dritt- und Viertklässler beim Problemlösen. Empirische Untersuchungen zur Typisierung

- spezifischer Problembearbeitungsstile. Berlin [u.a.]: LIT-Verl. (Begabungsforschung, 4).
- Germ, Michael; Harms, Ute (2010): Aufgabentypen und Anforderungsbereiche in Tests zur schriftlichen Leistungsmessung im Biologieunterricht. In: Druckausg.: Münster: Inst. f. d. Didaktik der Biologie der WWU, 2009 IDB Münster, Ber. Inst. Didaktik Biologie, Band 17 (2009), S. 1-17 ISSN 1868-7504.
- Glaser, Barney G.; Holton, Judith (2004): Remodeling Grounded Theory. In: Forum Qualitative Sozialforschung 5 (2).
- Goffman, Erving (1972): Asyl. Über die soziale Situation psychiatrischer Patienten und anderer Insassen. Frankfurt a. M.: Suhrkamp.
- Gogolin, Ingrid; Neumann, Ursula; Roth, Hans-Joachim (2003): Förderung von Kindern und Jugendlichen mit Migrationshintergrund. Gutachten. Bonn: Bundesländer-Kommission für Bildungsplanung und Forschungsförderung (Materialien zur Bildungsplanung und zur Forschungsförderung, 107).
- Gold, R. L. (1958): Roles in Sociological Field Observations. In: Social Forces 36 (3), S. 217–223.
- Graf, Dittmar (2001): Welche Aufgabentypen gibt es? In: Mathematisch naturwissenschaftlicher Unterricht 54 (7), S. 422–425.
- Gräsel, Cornelia; Parchmann, Ilka (2004): Implementationsforschung - oder: der steinige Weg, Unterricht zu verändern. In: Unterrichtswissenschaft 32 (3), S. 196–214.
- Gräsel, Cornelia; Parchmann, Ilka; Puhl, Thomas; Baer, Anja; Fey, Anja; Demuth, Reinhard (2004): Lehrerfortbildungen und ihre Wirkungen auf die Zusammenarbeit von Lehrkräften und die Unterrichtsqualität. In: Manfred Prenzel (Hg.): Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung. Münster: Waxmann.
- Gropengießer, Harald (Hg.) (2006): Mit Aufgaben lernen. Unterricht und Material 5-10. Unter Mitarbeit von Dietmar Höttecke, Telsche Nielsen und Lutz Stäudel: Friedrich Verlag.
- Grundgesetz für die Bundesrepublik Deutschland: "Grundgesetz für die Bundesrepublik Deutschland in der im Bundesgesetzblatt Teil III, Gliederungsnummer 100-1, veröffentlichten bereinigten Fassung, das zuletzt durch Artikel 1 des Gesetzes vom 23. Dezember 2014 (BGBl. I S. 2438) geändert worden ist". GG Art 3, Absatz 2.
- Günther, Christine (2013): Problemlösestrategien mathematisch begabter Kinder im Grundschulalter. In: Beiträge zum Mathematikunterricht 2013 Digital.

- Gurlitt, J.; Renkl, A. (2008): Are high-coherent concept maps better for prior knowledge activation? Differential effects of concept mapping tasks on high school vs. university students. In: *Journal of Computer Assisted Learning* 24 (5), S. 407–419.
- Hall, Richard H.; O'Donnell, Angela (1996): Cognitive and Affective Outcomes of Learning from Knowledge Maps. In: *Contemporary Educational Psychology* 21 (1), S. 94–101.
- Halpern, Diane F. (2007): *The science of sex differences in science and mathematics*. Malden: Blackwell Publ (Psychological science in the public interest, 8.2007,1, Suppl.).
- Halpern, Diane F. (2012): *Sex differences in cognitive abilities* (4th ed.), S. xiii, 453.
- Hammann, Marcus (2006): Kompetenzförderung und Aufgabenentwicklung. In: *Der mathematische und naturwissenschaftliche Unterricht* 59 (2), S. 85–95.
- Hampson, Elizabeth (1990): Variations in sex-related cognitive abilities across the menstrual cycle. In: *Brain and Cognition* 14 (1), S. 26–43.
- Hannover, Bettina (2008): Vom biologischen zum psychologischen Geschlecht: Die Entwicklung von Geschlechtsunterschieden (PSYNDEXshort), S. S 339-388.
- Hänze, Martin (2009): Mit Heterogenität umgehen. In: *Naturwissenschaften im Unterricht. Chemie* 20 (111/112), S. 2–4.
- Hänze, Martin; Schmidt-Weigand, Florian; Blum, Simone (2007): Mit gestuften Lernhilfen im naturwissenschaftlichen Unterricht lernen und arbeiten. In: Kerstin Rabenstein und u.a. (Hg.): *Kooperatives und selbstständiges Arbeiten von Schülern. Zur Qualitätsentwicklung von Unterricht*.
- Hartley, Bonny L.; Sutton, Robbie M. (2013): A stereotype threat account of boys' academic underachievement. In: *Child Development* 84 (5), S. 1716–1733.
- Hattie, John (2009): *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London, New York: Routledge.
- Haugwitz, Marion; Sandmann, Angela (2009): Kooperatives Concept Mapping in Biologie: Effekte auf den Wissenserwerb und die Behaltensleistung. In: *Zeitschrift für Didaktik der Naturwissenschaften* 15, 2009, S. 89–107, zuletzt geprüft am 25.11.2011.
- Hauser, Sabine; Nückles, Matthias; Renkl, Alexander (2006): Supporting concept mapping for learning from text. In: *Proceedings of the 7th international conference on Learning sciences*. Bloomington, Indiana: International Society of the Learning Sciences, S. 243–249.

- Hausmann, M.; Slabbekoorn, D.; van Goozen, S. H.; Cohen-Kettenis, P. T.; Gunturkun, O. (2000): Sex hormones affect spatial abilities during the menstrual cycle. In: *Behav Neurosci* 114 (6), S. 1245–1250.
- Häußler, P.; Lind, G. (2000): "Aufgabenkultur"- Was ist das? In: *Praxis der Naturwissenschaften* 49 (4), S. 2–10.
- Hay, David; Kinchin, Ian; Lygo-Baker, Simon (2008): Making learning visible: the role of concept mapping in higher education. In: *Studies in Higher Education* 33 (3), S. 295–311.
- Hayes, Andrew F.; Krippendorff, Klaus (2007): Answering the Call for a Standard Reliability Measure for Coding Data. In: *Communication Methods and Measures* 1 (1), S. 77–89.
- Heinze, Astrid (2005): Lösungsverhalten mathematisch begabter Grundschul Kinder - aufgezeigt an ausgewählten Problemstellungen. Münster: Lit (Begabungsforschung, 3).
- Helbig, Marcel (2012): Warum bekommen Jungen schlechtere Schulnoten als Mädchen? Ein sozialpsychologischer Erklärungsansatz. In: *Z f Bildungsforsch* 2 (1), S. 41-54.
- Heller, K. A. von; Perleth, C.: KFT 4-12+ R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision.
- Hellrung, Miriam (2009): Anforderungen im individualisierten Unterricht. In: *PÄD-Forum : unterrichten erziehen* (3), S. 113–115.
- Helmke, Andreas (2003): Unterrichtsqualität erfassen, bewerten, verbessern. Seelze: Kallmeyer.
- Helmke, Andreas (2006a): Was wissen wir über guten Unterricht? Über die Rückbesinnung auf den Unterricht als Kerngeschäft der Schule. In: *Pädagogik (Weinheim)* 58 (2), S. 42–45.
- Helmke, Andreas (2006b): Was wissen wir über guten Unterricht? Wissenschaftliche Erkenntnisse zur Unterrichtsforschung und Konsequenzen für die Unterrichtsentwicklung. Bertelsmann-Stiftung und Ministerium für Schule und Weiterbildung NRW. Essen, 2006, zuletzt geprüft am 14.06.2013.
- Helmke, Andreas (2007): Unterrichtsqualität erfassen, bewerten, verbessern. 5. Aufl. Seelze: Klett Kallmeyer (Schulisches Qualitätsmanagement).
- Hesse, Hermann-Günter (1994): Lehr-Lern-Zeit und Lernerfolg aus psychologischer Sicht. In: *Die Zeitdimension in der Schule als Gegenstand des Vergleichs*. Köln: Böhlau in Komm. (Studien und Dokumentationen zur vergleichenden Bildungsforschung, 61), S. 143–161.

- Hilbert, Tatjana S.; Renkl, Alexander (2008): Concept mapping as a follow-up strategy to learning from texts: what characterizes good and poor mappers? In: *Instr Sci* 36 (1), S. 53-73.
- Hinz, Andreas (2003): Inklusion - mehr als nur ein neues Wort!? Begriff der Inklusion und Abgrenzung zu Integration. In: *Lernende Schule* 6 (23), S. 15–17.
- Hoffmann, Lore (1997): An den Interessen von Jungen und Mädchen orientierter Physikunterricht. Ergebnisse eines BLK-Modellversuchs. Kiel: IPN.
- Holsanova, Jana; Holmberg, Nils; Holmqvist, Kenneth (2009): Reading information graphics: The role of spatial contiguity and dual attentional guidance. In: *Appl. Cognit. Psychol.* 23 (9), S. 1215–1226.
- Horz, Holger; Schnotz, W. (2010): Cognitive Load in Learning with Multiple Representations. In: Jan L. Plass, Roxana Moreno und Roland Brünken (Hg.): *Cognitive load theory*. Cambridge; New York: Cambridge University Press, S. 229–252.
- Huerta, Maria del Carmen; Adema, Willem; Baxter, Jennifer; Corak, Miles; Deding, Mette (2011): Early maternal employment and child development in five OECD countries. Paris (OECD social, employment and migration working papers).
- Jacobs, Bernhard (2008): Was wissen wir über die Lernwirksamkeit von Aufgabenstellungen und Feedback. In: Josef Thonhauser (Hg.): *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik*. Münster: Waxmann, S. 99–111.
- Jatzwauk, Paul (2007): *Aufgaben im Biologieunterricht. Eine Analyse der Merkmale und des didaktisch-methodischen Einsatzes von Aufgaben im Biologieunterricht*. Berlin: Logos-Verl.
- Jatzwauk, Paul; Rumann, Stefan; Sandmann, Angela (2008): Der Einfluss des Aufgabeneinsatzes im Biologieunterricht auf die Lernleistung der Schüler - Ergebnisse einer Videostudie. In: *Zeitschrift für Didaktik der Naturwissenschaften* 14, S. 263–283.
- Jatzwauk, Paul; Sandmann, Angela (2007): Analyse des Einsatzes von Aufgaben im Biologieunterricht der Klasse 9. In: Horst Bayrhuber (Hg.): *Ausbildung und Professionalisierung von Lehrkräften. Abstracts der Internationalen Tagung der Fachgruppe Biologiedidaktik im VBIO - Verband Biologie, Biowissenschaften & Biomedizin, 16.09 - 20.09.2007*. Kassel: Universität Kassel.
- Kalyuga, Slava (2007): Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction. In: *Educ Psychol Rev* 19 (4), S. 509-539.

- Kalyuga, Slava (2013): Enhancing transfer by learning generalized domain knowledge structures. In: *Eur J Psychol Educ* 28 (4), S. 1477–1493.
- Kalyuga, Slava; Ayres, Paul; Chandler, Paul; Sweller, John (2003): The Expertise Reversal Effect. *Educational Psychologist*. In: *Educational Psychologist* 38 (1), S. 23–31.
- Kamp, Rachele; Dolmans, Diana; Berkel, Henk; Schmidt, Henk (2012): The relationship between students' small group activities, time spent on self-study, and achievement. In: *Higher Education* 64 (3), S. 385–397.
- Kattmann, Ulrich (2003): Vom Blatt zum Planeten. Scientific literacy und kumulatives Lernen im Biologieunterricht und darüber hinaus. In: *PISA 2000 als Herausforderung*. Baltmannsweiler: Schneider-Verl. Hohengehren, S. 115–137.
- Kauertz, Alexander; Fischer, Hans E.; Mayer, Jürgen; Sumfleth, Elke; Walpuski, Maik (2010): Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. In: *Zeitschrift für Didaktik der Naturwissenschaften* 16.
- Killermann, Wilhelm; Hiering, Peter; Starosta, Bernhard (2005): *Biologieunterricht heute. Eine moderne Fachdidaktik*. 11. Aufl. Donauwörth: Auer.
- Kinchin, Ian M. (2000): Concept mapping in biology. In: *Journal of Biological Education (Society of Biology)* 34 (2), S. 61.
- Kinchin, Ian M.; Hay, David B.; Adams, Alan (2000): How a Qualitative Approach to Concept Map Analysis Can Be Used To Aid Learning by Illustrating Patterns of Conceptual Development. In: *Educational Research* 42 (1), S. 43–57.
- Klauer, Karl Josef (2010): *Transfer des Lernens. Warum wir oft mehr lernen als gelehrt wird*. Stuttgart: Kohlhammer Verlag.
- Klein, Michael; Pollmann, Anne (2007): *Die soziale Welt als Bühne. Ein kommunikationssoziologisches Forschungskolloquium*. Universität Hohenheim. Hohenheim, 2007.
- Klieme, Eckhard; Artelt, Cordula; Hartig, Johannes; Jude, Nina; Köller, Olaf; Prenzel, Manfred et al. (Hg.) (2010): *PISA 2009. Bilanz nach einem Jahrzehnt*. Münster, New York, NY, München, Berlin: Waxmann.
- Klieme, Eckhard; Rakoczy, Katrin (2008): Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. Paralleltitel: Empirical research on instruction and special didactics. In: *Zeitschrift für Pädagogik* 54 (2), S. 222–237.
- Klippert, Heinz (2008): *Methoden-Training. Bausteine für den Unterricht*. Sonderausg. Weinheim: Beltz.

- Klöpfel, Kathrin; Schwanewedel, Julia; Mayer, Jürgen (2013): Biologische Informationen aus unterschiedlichen Repräsentationen erschließen. Entwicklung und Überprüfung eines Testinstruments. In: Erkenntnisweg Biologiedidaktik (12), S. 27–42.
- KMK (2004): Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie. Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 05.02.2004. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Online verfügbar unter Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie, zuletzt geprüft am 07.08.2015.
- KMK (2005): Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004. München: Luchterhand.
- Knoll, Steffen (1998): Anforderungsgestaltung im Mathematikunterricht. In: Mathematik lehren, 1998 (90), S. 47–51.
- Koch-Priewe, Barbara; Niederbacher, Arne; Textor, Annette; Zimmermann, Peter (2009): Jungen - Sorgenkinder oder Sieger? Ergebnisse einer quantitativen Studie und ihre pädagogischen Implikationen. In: Jungen - Sorgenkinder oder Sieger?
- Koffka, Kurt (1999): Principles of Gestalt psychology. London: Routledge (International library of psychology, 7).
- Köhler, Karlheinz (2004): Welche fachgemäßen Arbeitsweisen werden im Biologieunterricht eingesetzt? In: Ulrike Spörhase-Eichmann und Wolfgang Ruppert (Hg.): Biologie-Didaktik. Praxishandbuch für die Sekundarstufe I und II. 1. Aufl. Berlin: Cornelsen Scriptor, S. 146–159.
- Köhler, Karlheinz (2010): Welche Medien werden im Biologieunterricht genutzt? In: Ulrike Spörhase-Eichmann (Hg.): Biologie: Methodik. Handbuch für die Sekundarstufe I und II. 1. Aufl. Berlin: Cornelsen Scriptor, S. 160–182.
- Köller, Olaf (2012): Mythen guter Unterrichtspraxis. In: IPN Blätter 29, 2012 (2), S. 1–3. Online verfügbar unter <http://www.ipn.uni-kiel.de/aktuell/ipnblatt/ip212/index.html>, zuletzt geprüft am 15.10.2012.
- Köller, Olaf; Trautwein, Ulrich; Lüdtke, Oliver; Baumert, Jürgen (2006): Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. In: Zeitschrift für Pädagogische Psychologie 20 (1-2), S. 27–39.
- Krammer, Kathrin (2009): Individuelle Lernunterstützung in Schülerarbeitsphasen. Eine videobasierte Analyse des Unterstützungsverhaltens von Lehrpersonen im Mathematikunterricht. Univ., Diss.--Zürich, 2008. Münster: Waxmann (Empirische Erziehungswissenschaft, 15).

- Krapp, Andreas (2006): Pädagogische Psychologie. Ein Lehrbuch. 5. Aufl. Weinheim [u.a.]: Beltz.
- Krippendorff, Klaus (2004): Content analysis. An introduction to its methodology. 2. Aufl. Thousand Oaks, Calif: Sage.
- Krippendorff, Klaus (2011): Computing Krippendorff's Alpha-Reliability. University of Pennsylvania. Philadelphia, Pennsylvania, USA. Online verfügbar unter http://repository.upenn.edu/cgi/viewcontent.cgi?article=1043&context=asc_papers, zuletzt geprüft am 31.03.2014.
- Krüger, Dirk; Meyfarth, Susanne (2009): Binnen - kurzer Zeit - differenzieren! In: Unterricht Biologie (347/348), S. 2–11.
- Kuckartz, Udo; Rädiker, Stefan; Ebert, Thomas; Schehl, Julia (2010): Statistik. Eine verständliche Einführung. In: Statistik.
- Kuhn, Jochen; Müller, Andreas (2007): Operationalisierung des Offenheitsgrades am Beispiel authentischer Aufgaben. In: Dietmar Höttecke (Hg.): Naturwissenschaftlicher Unterricht im internationalen Vergleich. [Gesellschaft für Didaktik der Chemie und Physik, Bd. 27 ; 33. Jahrestagung der GDGP vom 18. bis 21. September 2006 in den Räumen der Universität Bern]. Berlin: LIT-Verl.
- Kühn, Svenja Mareike (2011): Weiterentwicklung der Aufgabenkultur im naturwissenschaftlichen Unterricht der gymnasialen Oberstufe und im Abitur. In: Zeitschrift für Didaktik der Naturwissenschaften 17, S. 35–55.
- Kultusministerium Baden-Württemberg: Verordnung des Kultusministeriums über die Notenbildung, vom 05.05.1983. Fundstelle: §7 NVO Art. 1-3.
- Kunze, Ingrid (2010): Individuelle Förderung in der Schule. Wissenschaftliche Befunde - Möglichkeiten - Probleme. Universität Osnabrück. Bad Wildbad.
- Lachmayer, Simone; Nerdel, Claudia; Prectl, Helmut (2007): Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht. In: Zeitschrift für Didaktik der Naturwissenschaften 13 (p145-160; 200 KB), S. 145-160; 200 KB.
- Leibold, Kurt; Klautke, Siegfried (1999): Lerneffektivität des Einsatzes gegenständlicher Modelle in Biologieleistungskursen des Gymnasiums. In: Zeitschrift für Didaktik der Naturwissenschaften 5 (1), S. 3–23.
- Leisen, Josef (2005): Zur Arbeit mit Bildungsstandards. Lernaufgaben als Einstieg und Schlüssel. In: Der mathematische und naturwissenschaftliche Unterricht 58 (5), S. 306–308.

- Leisen, Josef (2006): Aufgabenkultur im mathematisch-naturwissenschaftlichen Unterricht. In: Der mathematische und naturwissenschaftliche Unterricht 59 (5), S. 260–266.
- Leisen, Josef (2010): Lernprozesse mithilfe von Lernaufgaben strukturieren. Informationen und Beispiele zu Lernaufgaben im kompetenzorientierten Unterricht. In: Naturwissenschaften im Unterricht. Physik 21, 2010 (117/118), S. 9–13.
- Leuders, Timo (2009): Neue Aufgabenkultur in der Mathematik. München: Oldenbourg.
- Leutner, Detlev; Fischer, Hans E.; Kauertz, Alexander; Schabram, Nina; Fischer, Jens (2008): Instruktionspsychologische und fachdidaktische Aspekte der Qualität von Lernaufgaben und Testaufgaben im Physikunterricht. In: Josef Thonhauser (Hg.): Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik. Münster: Waxmann, S. 169–182.
- Linden, Alexander (2007): Teilnehmende Beobachtung. Pädagogische Hochschule Freiburg. Online verfügbar unter <https://www.ph-freiburg.de/quasus/einstiegstexte/erhebungsinstrumente/beobachtungsverfahren/teilnehmende-beobachtung.html>, zuletzt geprüft am 07.08.2015.
- Lind, Gunter; Friege, Gunnar; Kleinschmidt, Lars; Sandmann, Angela (2004): Beispiellernen und Problemlösen. In: Zeitschrift für Didaktik der Naturwissenschaften 10, S. 29–49.
- Lipowsky, Frank; Rzejak, Daniela (2012): Lehrerinnen und Lehrer als Lerner - Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen. In: Schulpädagogik heute 3 (5), S. 1–17.
- Lüders, Christian (2006): Teilnehmende Beobachtung. In: Ralf Bohnsack (Hg.): Hauptbegriffe qualitativer Sozialforschung. 2. Aufl. Opladen u.a: Budrich (UTB, 8226), S. 151–153.
- Lupatsch, Judith; Hadjar, Andreas (2011): Determinanten des Geschlechterunterschieds im Schulerfolg: Ergebnisse einer quantitativen Studie aus Bern. In: Andreas Hadjar (Hg.): Geschlechtsspezifische Bildungsungleichheiten. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 177–202.
- Luthiger, Herbert (2008): An Aufgaben Professionalität entwickeln. In: Journal für LehrerInnenbildung : jlb. Aufgaben. Innsbruck: Studien-Verl (4), S. 37–46.
- Maaz, Kai; Baeriswyl, Franz; Trautwein, Ulrich (2011): Herkunft zensiert? Leistungsdiagnostik und soziale Ungleichheiten in der Schule ; eine Studie im

Auftrag der Vodafone Stiftung Deutschland. Düsseldorf: Vodafone Stiftung Deutschland.

- Mandl, Heinz; Kopp, Birgitta; Dvorak, Susanne (2004): Aktuelle theoretische Ansätze und empirische Befunde im Bereich der Lehr-Lern-Forschung. Schwerpunkt Erwachsenenbildung. Deutsches Institut für Erwachsenenbildung. Online verfügbar unter www.die-bonn.de/esprid/dokumente/doc-2004/mandl04_01.pdf, zuletzt geprüft am 28.07.2015.
- Markl, Jürgen; Sadava, David (2011): Purves Biologie. 9. Aufl. Heidelberg: Spektrum der Wissenschaft.
- Ma, Xin (2007): Gender Differences in Learning Outcomes. Paper commissioned for the EFA Global Monitoring Report 2008. Education for All by 2015: will we make it?
- Mayer, Jürgen; Wellnitz, Nicole; Klebba, Nicola; Kampa, Nele (2013): Kompetenzstufenmodelle für die naturwissenschaftlichen Fächer. In: Hans Anand Pant (Hg.): IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. Münster u.a: Waxmann, S. 74–83.
- Mayer, Richard E. (2001): Multimedia learning. 1. Aufl. Cambridge [u.a.]: Cambridge Univ. Press.
- Mayer, Richard E. (2004): Should There Be a Three-Strikes Rule against Pure Discovery Learning? In: American Psychologist 59 (1), S. 14–19.
- Mayer, Richard E.; Moreno, Roxana (2003): Nine Ways to Reduce Cognitive Load in Multimedia Learning. In: Educational Psychologist 38 (1), S. 43–52.
- McElvany, Nele; Schroeder, Sascha; Hachfeld, Axinja; Baumert, Jürgen; Richter, Tobias; Schnotz, Wolfgang et al. (2009): Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. In: Zeitschrift für Pädagogische Psychologie 23 (3-4), S. 223–235.
- Meisert, Anke (2012): Mit Modellen arbeiten. In: Steffen Schaal und Holger Weitzel (Hg.): Biologie unterrichten: planen, durchführen, reflektieren. 1. Aufl. Berlin: Cornelsen, S. 105–117.
- Meyer, Hilbert (2004): Was ist guter Unterricht? Berlin: Cornelsen.
- Müller, Andreas; Helmke, Andreas (2008): Qualität von Aufgaben als Merkmale der Unterrichtsqualität verdeutlicht am Fach Physik. In: Josef Thonhauser (Hg.): Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik. Münster: Waxmann, S. 31–46.

- Narciss, Susanne; Proske, Antje (2001): Lern- und Studieraufgaben für die universitäre Lehre. Unveröffentlichtes Manuskript. Technische Universität Dresden.
- Nesbit, John C.; Adesope, Olusola O. (2006): Learning with Concept and Knowledge Maps: A Meta-Analysis. In: *Review of Educational Research* 76 (3), S. 413–448.
- Neuhaus, Birgit; Vogt, Helmut (2005): Dimensionen zur Beschreibung verschiedener Biologielehrertypen auf Grundlage ihrer Einstellung zum Biologieunterricht. In: *Zeitschrift für Didaktik der Naturwissenschaften* 11, S. 73-84; 158 KB.
- Neumann, Knut; Kauertz, Alexander; Lau, Anna; Notarp, Hendrik; Fischer, Hans E. (2007): Die Modellierung physikalischer Kompetenz und ihrer Entwicklung. In: *Zeitschrift für Didaktik der Naturwissenschaften* 13, S. 101–121.
- Neuweg, Georg Hans (2008): Zur Funktion von Aufgaben im Lichte des tacit knowing view. In: Josef Thonhauser (Hg.): *Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik.* Münster: Waxmann, S. 83–96.
- Nido, Miriam; Trachsler, Ernst; Ackermann, Kurt; Brüggel, Susanne; Ulich, Eberhard (2008): *Arbeitsbedingungen, Belastungen und Ressourcen von Lehrpersonen und Schulleitungen im Kanton Aargau 2008. Ergebnisse der Untersuchung im Auftrag des Departements Bildung, Kultur und Sport (BKS, Kanton Aargau).* Aarau: BKS.
- Novak, J. D.; Cañas, A. J. (2006): *The Theory Underlying Concept Maps and How to Construct and Use Them.* Online verfügbar unter http://web.stanford.edu/dept/SUSE/projects/ireport/articles/concept_maps/The%20Theory%20Underlying%20Concept%20Maps.pdf, zuletzt geprüft am 07.08.2015.
- Nunan, David; Keobke, Ken (1995): Task Difficulty from the Learner's Perspective: Perceptions and Reality. In: *Hong Kong Papers in Linguistics and Language Teaching* 18, S. 1–12.
- O'Donnell, AngelaM; Dansereau, DonaldF; Hall, RichardH (2002): Knowledge Maps as Scaffolds for Cognitive Processing. In: *Educational Psychology Review* 14 (1), S. 71-86.
- OECD (2009): *OECD Factbook 2008: OECD Publishing*, zuletzt geprüft am 08.11.2013.
- OECD (2014): *PISA 2012 Results. Creative Problem Solving (Volume V): OECD Publishing.*

- OECD (2015): The ABC of gender equality in education. Aptitude, behaviour, confidence.
- Paivio, A. (1986): Mental representations: a dual coding approach. Oxford, England: Oxford University Press.
- Paradies, Liane; Linser, Hans Jürgen; Greving, Johannes (2007): Diagnostizieren, Fordern und Fördern. Berlin: Cornelsen Scriptor.
- Patterson, Michael E.; Dansereau, Donald F.; Wiegmann, Douglas A. (1993): Receiving information during a cooperative episode: Effects of communication aids and verbal ability. In: Learning and Individual Differences 5 (1), S. 1–11.
- Pietsch, Marcus (2010): Evaluation von Unterrichtsstandards. In: Zeitschrift für Erziehungswissenschaft (Zeitschrift für Erziehungswissenschaft) 13 (1), S. 121–148.
- Piopunuk, Marc; Wößmann, Ludger (2011): Folgekosten unzureichender Bildung durch entgangenes Wirtschaftswachstum. In: Wirtschaftsdienst 91 (S1), S. 34–41.
- Plant, E. Ashby; Ericsson, K. Anders; Hill, Len; Asberg, Kia (2005): Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. In: Contemporary Educational Psychology 30 (1), S. 96–116.
- Plass, Jan L.; Chun, Dorothy M.; Mayer, Richard E.; Leutner, Detlev (2003): Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. In: Computers in Human Behavior 19 (2), S. 221–243.
- Plotnick, Eric (1997): Concept mapping. A graphical system for understanding the relationship between concepts: an ERIC digest. Syracuse, NY: Clearinghouse on Information & Technology.
- Polly, Drew; Hannafin, Michael J. (2011): Examining How Learner-Centered Professional Development Influences Teachers' Espoused and Enacted Practices. In: The Journal of Educational Research 104 (2), S. 120–130.
- Porst, Rolf (2009): Fragebogen. Ein Arbeitsbuch. 2. Aufl. Wiesbaden: VS, Verl. für Sozialwiss. (Studienskripten zur Soziologie).
- Prändl, Ingeborg (2011): Die "Standard"-Fehler der Wahrnehmung. Online verfügbar unter <http://wahrnehmung.psychowissen.net/wahrnehmung/die-standard-fehler-der-wahrnehmung/index.html>, zuletzt geprüft am 08.04.2014.

- Prediger, Susanne (2008): Mit der Vielfalt rechnen. Aufgaben, Methoden und Strukturen für den Umgang mit Heterogenität im Mathematikunterricht, zuletzt geprüft am 02.11.2009.
- Prenzel, Manfred (2007): PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie. Münster, München [u.a.]: Waxmann, zuletzt geprüft am 17.06.2013.
- Prüfer, Peter; Stiegler, Angelika (2002): Die Durchführung standardisierter Interviews: Ein Leitfaden. In: ZUMA How-to-Reihe 11.
- Quenzel, Gudrun; Hurrelmann, Klaus (2010): Geschlecht und Schulerfolg: Ein soziales Stratifikationsmuster kehrt sich um. In: Köln Z Soziol 62 (1), S. 61-91.
- Rademann, Eva; Blume, Bernd (2001): Evaluierter Einsatz von Aufgaben im mathematisch-naturwissenschaftlichen Unterricht. In: Der mathematische und naturwissenschaftliche Unterricht 54 (7), S. 431–434.
- Rasch, Björn; Friese, Malte; Hofmann, Wilhelm Johann; Naumann, Ewald (2006a): Quantitative Methoden. Einführung in die Statistik. 2. Aufl. 2 Bände. Heidelberg: Springer (Quantitative Methoden, 1).
- Rasch, Björn; Friese, Malte; Hofmann, Wilhelm Johann; Naumann, Ewald (2006b): Quantitative Methoden. Einführung in die Statistik 2. 2. Aufl. 2 Bände. Heidelberg [u.a.]: Springer-Medizin-Verl. (2).
- Reißmüller, Ronny (2009): Methoden der qualitativen Sozialforschung. TU Chemnitz. Chemnitz, 2009.
- Reyer, Thomas; Trendel, Georg; Fischer, Hans E. (2004): Was kommt beim Schüler an? - Lehrerintentionen und Schülerlernen im Physikunterricht. In: Manfred Prenzel (Hg.): Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung. Münster: Waxmann, S. 195–211.
- Ruble, Diane N.; Martin, Carol Lynn (1998): Gender development, S. 933–1016.
- Ryan, Katherine E.; Ryan, Allison M. (2005): Psychological Processes Underlying Stereotype Threat and Standardized Math Test Performance. In: Educational Psychologist 40 (1), S. 53–63.
- Schaal, Steffen; Weitzel, Holger (Hg.) (2012): Biologie unterrichten: planen, durchführen, reflektieren. 1. Aufl. Berlin: Cornelsen.
- Schecker, Horst; Parchmann, Ilka (2006): Modellierung naturwissenschaftlicher Kompetenz. In: Zeitschrift für Didaktik der Naturwissenschaften 12, S. 45–66.

- Schmidt, Robert F.; Thews, Gerhard; Lang, Florian (2000): Physiologie des Menschen. Mit 96 Tabellen. 28. Aufl. Berlin [u.a.]: Springer (Springer-Lehrbuch).
- Schmiemann, Philipp (2012): Fachsprache in biologischen Testaufgaben. In: Zeitschrift für Didaktik der Naturwissenschaften 17, S. 115–136.
- Schmiemann, Philipp (2013): Aufgaben. In: Unterricht Biologie 37 (387/388), S. 2–9.
- Schneider, Christoph; Ludwig, Peter (2012): Auswirkungen von Maßnahmen der inneren Leistungsdifferenzierung auf Schulleistung und Fähigkeitsselbstkonzept im Vergleich zur äußeren Differenzierung. In: Beate Wischer (Hg.): Binnendifferenzierung. Teil 1: Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht. Unter Mitarbeit von Thorsten Bohl, Manfred Bönsch und Matthias Trautmann. neue Ausg. Immenhausen Hess: Prolog-Verlag (17), S. 72–106.
- Schnotz, Wolfgang; Bannert, Maria (1999): Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. In: Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie) 46 (3), S. 217–236. Online verfügbar unter <http://psycontent.metapress.com/content/11457ru4368211v8/fulltext.html>, zuletzt geprüft am 01.10.2012.
- Schnotz, Wolfgang; Bannert, Maria (2003): Construction and interference in learning from multiple representation. In: External and Internal Representations in Multimedia Learning 13 (2), S. 141–156.
- Scholz, Ingvelde (2008): Es ist normal, verschieden zu sein. Unterrichten in heterogenen Klassen. In: Ingvelde Scholz (Hg.): Der Spagat zwischen Fördern und Fordern. Unterrichten in heterogenen Klassen. Göttingen: Vandenhoeck & Ruprecht, S. 7–23.
- Scholz, Ingvelde (2012): Das heterogene Klassenzimmer. Differenziert unterrichten. Göttingen: Vandenhoeck & Ruprecht.
- Schöne, Helmar (2003): Die teilnehmende Beobachtung als Datenerhebungsmethode in der Politikwissenschaft. Methodologische Reflexion und Werkstattbericht. In: Forum Qualitative Sozialforschung / Forum: Qualitative Social Research; Vol 4, No 2 (2003): Subjectivity and Reflexivity in Qualitative Research II 4 (2).
- Schultheis, Klaudia (2012): Jungenforschung. Ergebnisse, Desiderate, Probleme. In: Bernd Ahrbeck und Michael Matzner (Hg.): Handbuch Jungen-Pädagogik. [aktualisiert und mit neuen Beiträgen]. 2. Aufl. Weinheim [u.a.]: Beltz (Pädagogik).
- Schwanewedel, Julia; Nitz, Sandra (2013): Repräsentationen im Biologieunterricht. Einflussfaktoren und Bedingungen für den Erwerb fachlicher

Kommunikationskompetenz. Internationale Tagung der Fachsektion Didaktik der Biologie (FDdB) im Vbio, 16.09.-20.09.2013. Universität Kassel, 2013.

- Seel, Norbert M. (1981): Lernaufgaben und Lernprozesse. Stuttgart: Kohlhammer (Studienbuch Pädagogik).
- Seufert, Tina; Brünken, Roland (2004): Supporting coherence formation in multimedia learning. In: Peter Gerjets, Paul Arthur Kirschner, J. Elen und R. Joiner (Hg.): Instructional design for effective and enjoyable computer-supported learning. Proceedings of the first joint meeting of the EARLI SIGs Instructional Design and Learning and Instruction with Computers. Tübingen: Knowledge Media Research Center, S. 138–147.
- Shulman, Lee S. (1986): Those Who Understand: Knowledge Growth in Teaching. In: Educational Researcher 15 (2), S. 4–14.
- Siebert, Horst (1998): Konstruktivismus. Konsequenzen für Bildungsmanagement und Seminargestaltung. Frankfurt/M: DIE (DIE / Deutsches Institut für Erwachsenenbildung (DIE), Pädagogische Arbeitsstelle des Deutschen Volkshochschul-Verbandes. Materialien für Erwachsenenbildung, 14).
- Sievert, Stephan; Kröhnert, Steffen (2015): Schwach im Abschluss. Warum Jungen in der Bildung hinter Mädchen zurückfallen - und was dagegen zu tun wäre. Berlin: Berlin-Institut für Bevölkerung und Entwicklung.
- Spörhase, Ulrike (Hg.) (2012a): Biologie-Didaktik. Praxishandbuch für die Sekundarstufe I und II. 5. Aufl. Berlin: Cornelsen.
- Spörhase, Ulrike (2012b): Welche Ziele verfolgt Biologieunterricht? In: Ulrike Spörhase (Hg.): Biologie-Didaktik. Praxishandbuch für die Sekundarstufe I und II. 5. Aufl. Berlin: Cornelsen, S. 25–67.
- Stanat, Petra; Rauch, Dominique; Segeritz, Michael (2010): Schülerinnen und Schüler mit Migrationshintergrund. In: Eckhard Klieme, Cordula Artelt, Johannes Hartig, Nina Jude, Olaf Köller, Manfred Prenzel et al. (Hg.): PISA 2009. Bilanz nach einem Jahrzehnt. Münster, New York, NY, München, Berlin: Waxmann.
- Stäudel, Lutz (2004): Aufgaben für den Chemieunterricht. Eine Einführung. In: Naturwissenschaften im Unterricht. Chemie 15 (82-83), S. 4–6.
- Stäudel, Lutz (2008): Aufgaben mit gestuften Hilfen für den Biologie-Unterricht. Aufgabensammlung mit Lernhilfen in unterschiedlichen Schwierigkeitsgraden. 1. Aufl. Seelze: Friedrich.
- Stäudel, Lutz (2009a): Aufgaben mit gestuften Hilfen. Eine selbstdifferenzierende Lernumgebung am Beispiel von Osmose und Verbrennung. In: Naturwissenschaften im Unterricht. Chemie 20 (111/112), S. 72–77.

- Stäudel, Lutz (2009b): Differenzieren im Chemieunterricht. Eine Herausforderung für Lehrkräfte, Lernende und das Selbstverständnis von Schule. In: Naturwissenschaften im Unterricht. Chemie 20 (111/112), S. 8–11.
- Steele, Claude M. (1997): A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. In: American Psychologist 52 (6), S. 613–629.
- Steele, Claude M.; Spencer, Steven J.; Aronson, Joshua (2002): Contending with group image: The psychology of stereotype and social identity threat. In: Mark P. Zanna (Hg.): Advances in Experimental Social Psychology: Academic Press, S. 379-440, x, 455.
- Stern, Elsbeth; Aprea, Carmela; Ebner, Herman G. (2003): Improving Cross-Content Transfer in Text Processing by Means of Active Graphical Representation. In: Learning and Instruction 13 (2), S. 191–203.
- Strunz-Maireder, Edith (2009): "Pedagogical content knowledge". Wie die Lehrer/innenwissenschaftler die Bedeutung der Fachkompetenz neu entdeckt. In: Wissen plus (5), S. 41–44.
- Studienseminar Koblenz (2011): Subjektive Fehlerquellen im Überblick. Studienseminar Koblenz. Koblenz, zuletzt geprüft am 08.04.2014.
- Sweller, John; van Merriënboer, Jeroen J. G.; Paas, Fred G. W. C. (1998): Cognitive Architecture and Instructional Design. In: Educational Psychology Review 10 (3), S. 251–296.
- Terzer, Eva; Upmeyer zu Belzen, Annette (2007): Naturwissenschaftliche Erkenntnisgewinnung durch Modelle. Modellverständnis als Grundlage für Modellkompetenz. In: Berichte des Institutes für Didaktik der Biologie der Westfälischen Wilhelms-Universität Münster, IDB (16), S. 33–56.
- Thonhauser, Josef: Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung allgemeiner Didaktik und Fachdidaktik /// Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik.
- Thonhauser, Josef (2008): Warum (neues) Interesse am Thema Aufgaben? In: Josef Thonhauser (Hg.): Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik. Münster: Waxmann, S. 13–25.
- Tillmann, Klaus-Jürgen (2004): System jagt Fiktion. Die homogene Lerngruppe. In: Friedrich-Jahresheft (22), S. 144 S.

- Tomlinson, Carol Ann; Brighton, Catherine; Hertberg, Holly; Callahan, Carolyn M.; Moon, Tonya R.; Brimijoin, Kay et al. (2003): Differentiating Instruction in Response to Student Readiness, Interest, and Learning Profile in Academically Diverse Classrooms: A Review of Literature. In: *Journal for the Education of the Gifted* 27 (2), S. 119–145.
- Toth, Eva Erdosne; Suthers, Daniel D.; Lesgold, Alan M. (2002): “Mapping to know”: The effects of representational guidance and reflective assessment on scientific inquiry. In: *Sci. Ed.* 86 (2), S. 264–286.
- Tsien, Joe Z. (2007): The Memory Code. (Cover story). In: *Scientific American* 297 (1), S. 52–59.
- Tzuriel, David; Egozi, Gila (2010): Gender Differences in Spatial Ability of Young Children: The Effects of Training and Processing Strategies. In: *Child Development* 81 (5), S. 1417–1430.
- Upmeier zu Belzen, Annette; Krüger, Dirk (2010): Modellkompetenz im Biologieunterricht. In: *Zeitschrift für Didaktik der Naturwissenschaften* 16 (p41-57; 0), S. 41-57; 0.
- van den Hurk, M. M.; Wolfhagen, H. A. P.; Dolmans, D. H. J. M.; van der Vleuten, C. P. M. (1998): The Relation Between Time Spent on Individual Study and Academic Achievement in a Problem-Based Curriculum. In: *Adv Health Sci Educ Theory Pract* 3 (1), S. 43-49.
- van Zele, Els; Lenaerts, Josephina; Wieme, Willem (2004): Improving the Usefulness of Concept Maps as a Research Tool for Science Education. In: *International Journal of Science Education* 26 (9), S. 1043–1064.
- VG Braunschweig, Urteil vom 01.09.2010, Aktenzeichen 6 B 182/10. In: *openJur* 2012, 51025.
- Volkholz, Sybille; Baumert, Jürgen; Fratton, Peter; Gräsel, Cornelia; Kunter, Mareike; Lemmermöhle, Doris et al. (2013): Expertenkommission zur Weiterentwicklung der Lehrerbildung in Baden-Württemberg. Hg. v. Forschung und Kunst Baden-Württemberg Ministerium für Wissenschaft. Stuttgart.
- Wai, Jonathan; Cacchio, Megan; Putallaz, Martha; Makel, Matthew C. (2010): Sex differences in the right tail of cognitive abilities: A 30 year examination. In: *Intelligence* 38 (4), S. 412–423. Online verfügbar unter <http://dx.doi.org/10.1016/j.intell.2010.04.006>.
- Walgenbach, Katharina (2013): Heterogenität als Chance. Revitalisierung der Differenzdebatte in der Schulpädagogik? In: *Schulpädagogik heute* 4 (8).
- Warrington, M.; Younger, M.; Williams, J. (2000): Student Attitudes, Image and the Gender Gap. In: *British Educational Research Journal* 26 (3), S. 393–407.

- Weidenmann, Bernd (2006): Lernen mit Medien. In: Pädagogische Psychologie, S. 423–476.
- Weinert, Franz E. (2001): Leistungsmessungen in Schulen. Weinheim: Beltz.
- Weinert, Franz Emanuel; Schrader, Friedrich-Wilhelm (1986): Diagnose des Lehrers als Diagnostiker. In: Karlheinz Petillon Hanns Wagner Jürgen W. L. Wolf Bernhard Ingenkamp (Hg.): Schülergerechte Diagnose. Theoretische und empirische Beiträge zur pädagogischen Diagnostik : Festschrift zum 60. Geburtstag von Karlheinz Ingenkamp. Weinheim: Beltz, S. 11–29.
- Weis, Mirjam; Heikamp, Tobias; Trommsdorff, Gisela (2013): Gender differences in school achievement. The role of self-regulation. Konstanz: Bibliothek der Universität Konstanz.
- Weitzel, Holger (2012): Aufgaben entwickeln und einsetzen. In: Steffen Schaal und Holger Weitzel (Hg.): Biologie unterrichten: planen, durchführen, reflektieren. 1. Aufl. Berlin: Cornelsen, S. 132–148.
- Wellnitz, Nicole; Fischer, Hans E.; Kauertz, Alexander; Mayer, Jürgen; Neumann, Irene; Pant, Hans Anand et al. (2012): Evaluation der Bildungsstandards. eine fächerübergreifende Testkonzeption für den Kompetenzbereich Erkenntnisgewinnung. In: Zeitschrift für Didaktik der Naturwissenschaften 18, S. 261–291.
- Wellnitz, Nicole; Mayer, Jürgen (2008): Evaluation von Kompetenzstruktur und -niveaus zum Beobachten, Vergleichen, Ordnen und Experimentieren. In: Kai Niebert (Hg.): Erkenntnisweg Biologiedidaktik 7. Unter Mitarbeit von Dirk Krüger, Annette Upmeier zu Belzen und Tanja Riemeier. Beiträge auf der 10. Frühjahrsschule der Sektion Biologiedidaktik im VBiO. Hannover, S. 129–143.
- White, H. B. (2002): How to construct a concept map. Department of Chemistry and Biochemistry, University of Delaware. Online verfügbar unter <http://www.udel.edu/chem/white/teaching/ConceptMap.html>, zuletzt geprüft am 27.02.2014.
- Wiegmann, Douglas A.; Dansereau, Donald F.; McCagg, Edward C.; Rewey, Kirsten L.; Pitre, Urvashi (1992): Effects of knowledge map characteristics on information processing. In: Contemporary Educational Psychology 17 (2), S. 136–155.
- Williamson, David F.; Parker, Robert A.; Kendrick, Juliette S. (1989): The Box Plot: A Simple Visual Method to Interpret Data. In: Annals of Internal Medicine 110 (11), S. 916–921.
- Winter, Felix (2008): Mit Aufgaben das Lernen sondieren. In: Josef Thonhauser (Hg.): Aufgaben als Katalysatoren von Lernprozessen. Eine zentrale Komponente

- organisierten Lehrens und Lernens aus der Sicht von Lernforschung, Allgemeiner Didaktik und Fachdidaktik. Münster: Waxmann, S. 115–129.
- Wirtz, Markus (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen: Hogrefe.
- Wischer, Beate (Hg.) (2012): Binnendifferenzierung. Teil 1: Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht. Unter Mitarbeit von Thorsten Bohl, Manfred Bönsch und Matthias Trautmann. neue Ausg. Immenhausen Hess: Prolog-Verlag (17).
- Wittrock, Merlin C. (1989): Generative Processes of Comprehension. *Educational Psychologist*. In: *Educational Psychologist* 24 (4), S. 345–376.
- Wodzinski, Rita (2007): Varianten. Aufgaben für Differenzierung umarbeiten. In: Rita Wodzinski, Christoph Wodzinski und Ralph Hepp (Hg.): *Differenzierung. Naturwissenschaften im Unterricht Physik* 18 (99/100): Erhard Friedrich Verlag, S. 16–21.
- Wodzinski, Rita; Wodzinski, Christoph (2007): Unterschiede zwischen Schülern - Unterschiede im Unterricht? In: Rita Wodzinski, Christoph Wodzinski und Ralph Hepp (Hg.): *Differenzierung. Naturwissenschaften im Unterricht Physik* 18 (99/100): Erhard Friedrich Verlag, S. 4–9.
- Woessmann, Ludger; Piopiunik, Marc (2009): Was unzureichende Bildung kostet. Eine Berechnung der Folgekosten durch entgangenes Wirtschaftswachstum ; [wirksame Bildungsinvestitionen]. Gütersloh: Bertelsmann Stiftung.
- Yildirim, Barış O.; Derksen, Jan J.L (2012): A review on the relationship between testosterone and life-course persistent antisocial behavior. In: *Psychiatry Research* 200 (2-3), S. 984–1010.
- Zimmermann, Gunter E. (2006): Methoden der empirischen Sozialforschung. In: Bernhard Schäfers und u.a. (Hg.): *Grundbegriffe der Soziologie*. 9., grundlegend überarb. und aktualisierte Aufl. Wiesbaden: VS Verl. für Sozialwissenschaften (Lehrbuch), S. 179–190.

9 ANHANG

9.1 Darstellung der zusätzlich eingesetzten Materialien

Um die Leistungen von Schülern beim Lösen von Aufgaben nachvollziehen zu können, müssen mindestens drei Bestandteile klar sein (Narciss und Proske 2001). Diese Bestandteile umfassen (1) die Klärung der Darbietung der nötigen Inhalte über entsprechende Repräsentationen, (2) die Darstellung der Aufgaben mit den erforderlichen kognitiven Operationen, und (3) die Klärung des Erwartungshorizonts der Aufgaben. Diese drei Bestandteile werden im Folgenden zuerst für die eingesetzten Repräsentationen beim Input und später für die Lernaufgaben dargestellt.

Wie bereits erwähnt so ist das Ziel einer Präsentation von Inhalten, die relevanten Bestandteile eines Themas darzustellen. Dies geschieht über Repräsentationsformen wie zum Beispiel Tabellen, Graphiken oder Abbildungen. In dieser Studie wurden, entsprechend den Forschungsfragen, die aufgabenrelevanten Inhalte primär über Text oder Experten-Concept-Map vermittelt und zusätzlich durch zwei weitere Repräsentationsformen ergänzt. Hierbei handelte es sich zum einen um ergänzende Abbildungen und zum anderen um ein Modell vom menschlichen Auge. Der Ablauf bei allen Interviews war identisch. Die Repräsentationsformen wurden den Lernenden nacheinander präsentiert und anschließend nach Schwierigkeiten beim Verständnis gefragt. Zuerst mussten sich die Lernenden alleine mit der jeweiligen Repräsentationsform auseinandersetzen. Im weiteren Lauf des Interviews sollten sie jedoch zunehmend mehr mit dem Partner diskutieren und die Repräsentationen dazu nutzen, um die Lernaufgaben in Partnerarbeit zu lösen.

Im Folgenden werden die eingesetzten Materialien dieser Studie vorgestellt. Hierbei wird sowohl auf deren Entwicklung als auch auf deren Auswertung eingegangen. Dadurch wird das Nachvollziehen und Interpretieren der Werte im Ergebnisteil transparent und selbsterklären. Die Instrumente werden entsprechend des Ablaufs der Studie in chronologischer Reihenfolge aufgeführt. Der Ablauf der Interviews ist in Tab. 41 auf der folgenden Seite dargestellt.

Tab. 41: Ablauf der Interviews mit Einsatz der Materialien.

Vor dem Interview	siehe Seite
• Einholen des Einverständnisses der Eltern	
• Schriftlicher Fragebogen zur Einschätzung der Schüler in <u>Einzelarbeit</u>	219
Während dem Interview	
• Empfang, Begrüßung und mündliches Einverständnis der Schüler	
• Auseinandersetzung mit dem Input in <u>Einzelarbeit</u>	
○ Der Fachtext ohne Farbcodierung	110
○ Der Fachtext mit Farbcodierung	221
○ Experten-Concept-Map ohne Farbcodierung	110
○ Experten-Concept-Map mit Farbcodierung	223
• Ausfüllen des Multiple-Choice-Tests zum Input in <u>Einzelarbeit</u>	223
• Selbstkontrolle der Ergebnisse durch die Schüler	
• Auseinandersetzung mit den Abbildungen in <u>Einzelarbeit</u>	231
• Ausfüllen des Multiple-Choice-Tests zu den Abbildungen in <u>Einzelarbeit</u>	231
• Selbstkontrolle der Ergebnisse durch die Schüler	
• Einführung und Vorstellung des Augenmodells	242
• Bearbeitung von Aufgabe 1 in <u>Partnerarbeit</u> (mit Hilfen zu Aufgabe 1)	112 (242)
• Bearbeitung von Aufgabe 2 in <u>Partnerarbeit</u> (mit Hilfen zu Aufgabe 2)	112 (244)
• Bearbeitung von Aufgabe 3 in <u>Partnerarbeit</u> (mit Hilfen zu Aufgabe 3)	112 (245)
• Bearbeitung von Aufgabe 4 in <u>Partnerarbeit</u> (mit Hilfen zu Aufgabe 4)	112 (246)
• Einschätzung der Aufgaben 1 bis 4 durch die Schüler	
• Danksagung an die Schüler und Verabschiedung	

9.1.1 Input: Fachtext

Der Aufbau der Netzhaut

Lese den Text durch und verschaffe dir einen Überblick über den Aufbau der Netzhaut.

Im hinteren Teil des **Auges** befindet sich die **Netzhaut**. In der Netzhaut des Menschen befinden sich rund 130 Millionen lichtempfindliche **Sehsinneszellen**. Die Sehsinneszellen sind die einzigen Zellen im Körper, die Lichtreize wahrnehmen können. Sie sind über andere Nervenzellen mit Schaltzellen verbunden. Die Fortsätze der **Schaltzellen** ziehen in den **Sehnerv**. Der Sehnerv führt zum **Gehirn**. An der Stelle, an der der Sehnerv den Augapfel verlässt, liegt der **blinde Fleck**. Im Bereich des blinden Flecks sind keine Sehsinneszellen vorhanden.

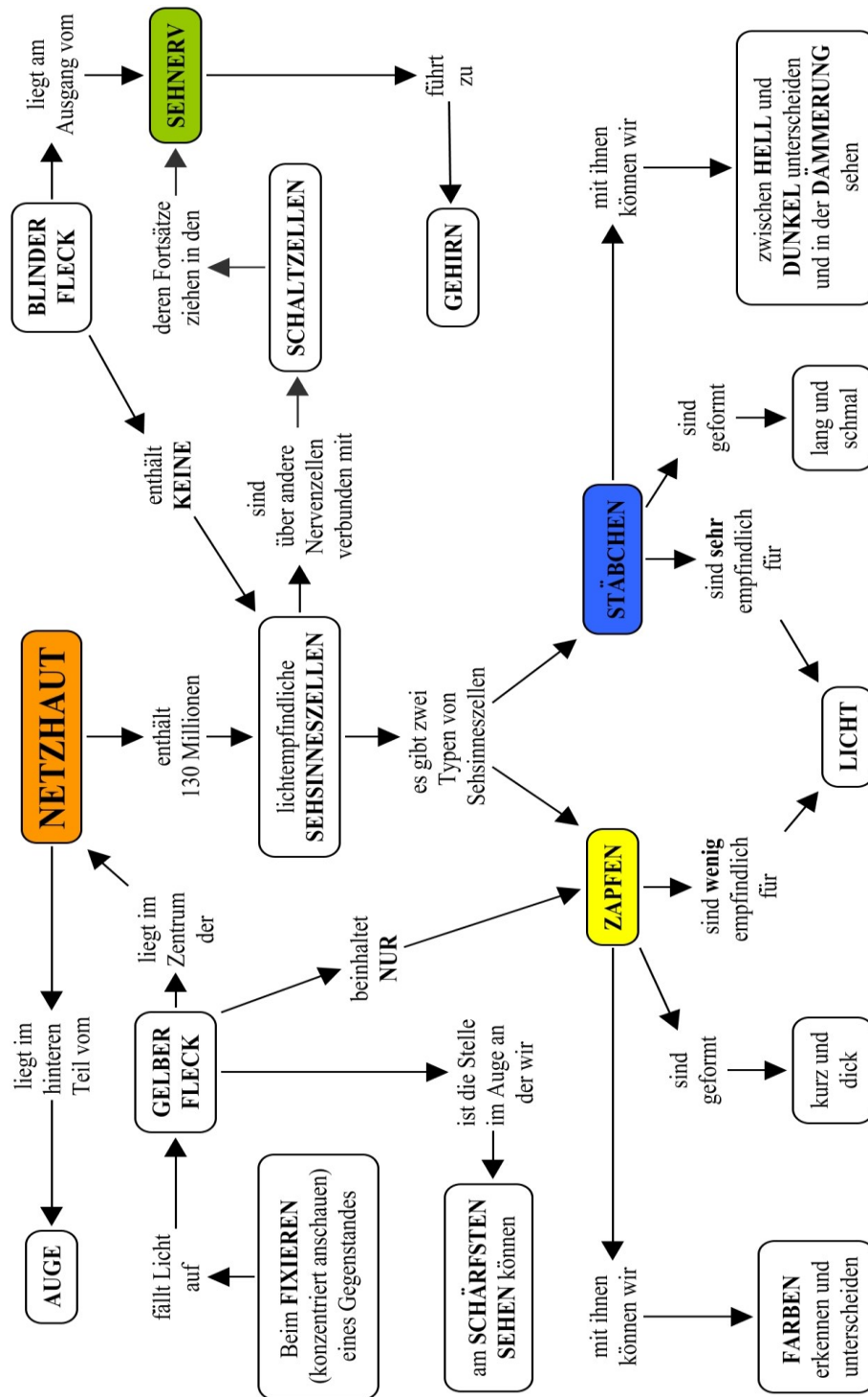
Es gibt zwei Typen von Sehsinneszellen, die Stäbchen und die Zapfen. Die **Stäbchen** sind **sehr** lichtempfindlich, mit ihnen können wir in der Dämmerung und nachts sehen. Allerdings können sie nur zwischen Hell und Dunkel unterscheiden. Die äußere Form der Stäbchen ist länglich und schmal. Die **Zapfen** sind **wenig** lichtempfindlich. Im Vergleich zu den Stäbchen sind die Zapfen in ihrer Form eher kurz und dick. Mit den Zapfen können wir allerdings Farben erkennen und unterscheiden.

Die Netzhaut ist nicht an allen Stellen gleich aufgebaut. Im Randbereich der Netzhaut gibt es nur Stäbchen. Im zentraleren Bereich befinden sich Stäbchen und Zapfen. An einer Stelle im Zentrum befinden sich **nur** Zapfen in der Netzhaut. Hier ist jeder Zapfen über andere Nervenzellen mit einer Schaltzelle verbunden. Daher liegt dort die Stelle des **schärfsten Sehens**. Diese Stelle wird der **gelbe Fleck** genannt. Beim **Fixieren** (konzentriert anschauen) eines Gegenstandes wird das Auge immer so ausgerichtet, dass das Bild direkt auf den gelben Fleck in der Netzhaut trifft.

9.1.2 Input: Experten-Concept-Map

Der Aufbau der Netzhaut

Gehe die Experten-Concept-Map Schritt für Schritt durch. Beginne bei dem Begriff *Netzhaut*. Wenn du alle Begriffe und Verbindungen gelesen hast, kannst du mit der nächsten Aufgabe beginnen.



9.1.3 Multiple-Choice-Test zu dem Input

Der Aufbau der Netzhaut

Lese dir zuerst alle Fragen durch und kreuze danach die richtigen Antworten an.

(**Achtung:** Manchmal sind mehrere Antworten richtig!)

Wo befindet sich die NETZHAUT im Auge?	a) Überall im Auge
	b) Im vorderen Teil des Auges
	c) Im hinteren Teil des Auges
	d) In der Mitte des Auges
Was wisst ihr über die STÄBCHEN?	a) Mit ihrer Hilfe kann man Farben erkennen und unterscheiden
	b) Mit ihnen kann man zwischen hell und dunkel unterscheiden
	c) Sie helfen beim Sehen bei Nacht und bei Dämmerung
	d) Sie sind direkt mit dem Gehirn verschaltet
	e) Sie sind eher lang und schmal geformt
	f) Sie sind eher kurz und dick geformt
	g) Sie benötigen viel Licht
	h) Sie benötigen wenig Licht
Was wisst ihr über die ZAPFEN	a) Mit ihrer Hilfe kann man Farben erkennen und unterscheiden
	b) Mit ihnen kann man zwischen hell und dunkel unterscheiden
	c) Sie helfen beim Sehen bei Nacht und bei Dämmerung
	d) Sie sind direkt mit dem Gehirn verschaltet
	e) Sie sind eher lang und schmal geformt
	f) Sie sind eher kurz und dick geformt
	g) Sie benötigen viel Licht
	h) Sie benötigen wenig Licht
	i) Mit ihrer Hilfe kann man Farben erkennen und unterscheiden
Was wisst ihr über den GELBEN FLECK?	a) Hier befinden sich keine Lichtsinneszellen
	b) Dies ist die Stelle des schärfsten Sehens
	c) Hier befinden sich nur Zapfen
	d) Hier befinden sich nur Stäbchen
	e) Hier befindet sich der Sehnerv
	f) Wenn man direkt etwas anschaut (fixiert), fällt das Licht direkt auf diese Stelle im Auge
Was wisst ihr über den BLINDEN FLECK?	a) Hier befinden sich keine Lichtsinneszellen
	b) Dies ist die Stelle des schärfsten Sehens
	c) Hier befinden sich nur Zapfen
	d) Hier befinden sich nur Stäbchen
	e) Hier befindet sich der Sehnerv
	f) Wenn man direkt etwas anschaut (fixiert), fällt das Licht direkt auf diese Stelle im Auge

9.1.4 Input: Abbildungen zum Auge

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

9.1.5 Multiple-Choice-Test zu den Abbildungen

Um zu gewährleisten, dass sich auch alle Schüler mit den Abbildungen auseinandersetzen, wurden die Schüler aufgefordert, die Abbildungen sorgfältig zu betrachten. Anschließend wurde ihr Verständnis überprüft. Dies geschah auf gleiche Weise wie zuvor beim Input. Zuerst sollten die Schüler sich in Einzelarbeit die Abbildungen anschauen und danach einen Multiple-Choice-Test ausfüllen. Der Multiple-Choice-Test enthielt fünf mögliche Titel für jede Abbildung. Die Schüler mussten sich für eine passende Überschrift entscheiden und diese ankreuzen. Ihr Ergebnis durften sie direkt im Anschluss selbst überprüfen und gegebenenfalls Verständnisfragen zu den Abbildungen stellen. Die Auswertungsmethode und Ergebnisse des Multiple-Choice-Tests befinden sich ab Seite 167, der original Fragebogen auf der folgenden Seite.

Der Aufbau der Netzhaut

Betrachte die Abbildungen. Du wirst feststellen, dass sich die farbigen Wörter in den Abbildungen wiederfinden. Achte bei den Abbildungen besonders auf Unterschiede im Aufbau der Netzhaut.

Abbildungen (siehe extra Blatt)

Entscheide dich, welche der vorgegebenen Überschriften du für die jeweilige Abbildung wählen würdest und kreuze diese an.

Wenn du fertig bist, kannst du deine Antworten mit deinem Partner vergleichen. Danach bekommst du die Lösung und kannst deine Antworten kontrollieren.

<p>Überschrift für Abbildung 1:</p> <ul style="list-style-type: none">a) Wo ist das Auge?b) Wir nehmen das Auge genauer unter die Lupec) Das Auge von innend) Das Auge des Menschene) Schnitt durch das Auge	<p>Überschrift für Abbildung 2:</p> <ul style="list-style-type: none">a) Funktion der Netzhautb) Das Auge von vornec) Das Auge wird näher betrachtetd) Der gelbe Fleck vom Augee) Schnitt durch das Auge
<p>Überschrift für Abbildung 4:</p> <ul style="list-style-type: none">a) Scharfer Sehbereichb) Der Sehnervc) Zapfen und Stäbchend) Der gelbe Flecke) Die Stäbchen der Netzhaut	<p>Überschrift für Abbildung 5:</p> <ul style="list-style-type: none">a) Der gelbe Fleckb) Der Sehnervc) Ort des schärfsten Sehensd) Der blinde Flecke) Zapfen und Stäbchen

9.2 Darstellung der kognitiven Operationen durch die Materialien

9.2.1 Fragebogen zur Einschätzung der Schüler

Wie bereits im „Abschnitt Rekrutierung und Beschreibung der Testpersonen“ (Seite 105 ff.) erwähnt wurde, diente der Fragebogen primär der Einteilung der Testpersonen in leistungsstärkere und leistungsschwächere Schüler. Ferner wurde dadurch die Eignung der Personen für die Studie durch entsprechende Ein- und Ausschlusskriterien überprüft. Der Fragebogen wurde so konzipiert, dass er von den beteiligten Lehrkräften ohne zusätzlichen Aufwand innerhalb von 25 Minuten im Unterricht eingesetzt werden konnte. Er wurde als Klassensatz an die Lehrkräfte verschickt und immer von der ganzen Klasse zeitgleich bearbeitet. Im Anschluss daran haben die Lehrkräfte die Fragebögen entsprechend der Einverständniserklärungen sortiert, die der teilnehmenden Schüler zurückgeschickt und die restlichen verworfen. Der Fragebogen bestand aus insgesamt sieben Aufgaben, auf die im Folgenden eingegangen wird.

Bei Aufgabe 1 wurde das Interesse der Schüler am Fach Biologie abgefragt. Hierfür sollten die Schüler sieben unterschiedliche Aussagen über Interesse an Biologie lesen und entscheiden inwieweit die Aussagen²³ auf sie persönlich zutreffen. Den Aussagen wurden bei der Auswertung der fünfstufigen Likert-Skala die Zahlen 1 bis 5 zugeordnet. Dadurch konnte folgende Aussage abgeleitet werden: Je höher der Wert, desto höher das Fachinteresse der betreffenden Person. Diese Aufgabe hatte keine Auswirkung auf die Auswahl der Schüler, sondern galt einzig und allein dem Forschungsinteresse.

Mit Aufgabe 2 wurde mögliches Vorwissen der Schüler zum Thema Auge abgefragt. Die Aufgabe bestand aus zwei geschlossenen und einer offenen Frage. Die beiden geschlossenen Fragen sollten klären, ob das Thema schon in der Schule behandelt wurde oder ob sich die Schüler bereits auf anderem Wege wie zum Beispiel durch Film, Fernsehen oder Internet, fundiertes Wissen zum Thema Auge angeeignet haben. Bei der dritten und offenen Frage sollten sie dieses Wissen in Stichworten niederschreiben. Zeigten Schüler hierbei mehr als drei richtige Antworten, wurde sie von der Studie ausgeschlossen, da zu viel Vorwissen die Ergebnisse der Studie verfälschen würden.

In Aufgabe 2 sollten die Schüler eine Experten-Concept-Map lesen und anschließend vier Multiple-Choice-Fragen beantworten. Die Frage hatte das Ziel zu prüfen, ob die Schüler über ein ausreichendes Verständnis für Experten-Concept-Maps verfügen oder nicht. Die Überprüfung war nötig, da solch ein Verständnis für die Studie von immenser Wichtigkeit war. Aus diesem Grund lag der Anspruch sehr hoch und so

²³ Die Auswahlmöglichkeiten lauteten: *stimmt*, *stimmt manchmal*, *kann ich nicht sagen*, *stimmt eher nicht*, *stimmt nicht*.

wurden alle Schüler, die weniger als 75% der Fragen richtig beantwortet hatten, von einem Interview mit Experten-Concept-Map als Input ausgeschlossen.

Neben dem Fachinteresse wurde in Aufgabe 4 die bevorzugte Sinnesmodalität (auditiv, visuell, haptisch) durch einen Lerntypentest nach Klippert erhoben (Klippert 2008). Analog zur Erhebung des Fachinteresses sollten die Schüler 17 Aussagen über ihr Lernverhalten lesen und durch ankreuzen bewerten, inwieweit diese Aussagen auf sie zutreffen. Die Auswahlmöglichkeiten lauteten: *so kann ich mir VIEL merken*, *so kann ich mir EINIGES merken* und *so kann ich mir WENIG merken*. Die Angaben der Schüler wurden entsprechend des Lerntyps in drei Kategorien aufgeteilt und mit Punkten versehen. Die Ergebnisse sind folgendermaßen zu deuten: Je höher die erreichte Punktezahl, desto ausgeprägter die jeweilige Sinnesmodalität. Aufgabe 4 hatte keine Auswirkung auf die Auswahl der Schüler.

Aufgabe 5 hatte die gleiche Funktion wie Aufgabe 2. Auch hier hatten die Schüler die Chance, ihr Vorwissen zum Thema Auge zu äußern. Die Frage der fünften Aufgabe war im Vergleich zu Aufgabe 2 jedoch spezifischer gehalten und richtete sich explizit auf den Aufbau vom menschlichen Auge. Die Bewertung der Antworten erfolgte, wie zuvor auch bei Aufgabe 2. Es wurden diejenigen Schüler von der Studie ausgeschlossen, die mehr als drei richtige Antworten gaben.

Bei den Aufgaben 6 und 7 sollte geprüft werden, ob die Schüler über ein ausreichendes Sprach- und Textverständnis verfügen. Um das feststellen zu können wurden Auszüge des kognitiven Fähigkeitstests (KFT 4-12+R) verwendet. Hierbei handelte es sich um die Subskalen von Test 1 für die Klassen 6 – 8 und Test 2 für die Klassen 5 – 8. Fielen Schüler bei dieser Aufgabe unter das Skalenniveau ihres Schultyps oder ihrer Klassenstufe, so wurden sie ebenfalls von dem Interview mit Text als Input ausgeschlossen. Dies war erforderlich, da ein unzureichendes Textverständnis den Text-Input des Interviews obsolet gemacht hätte und so das Testergebnis verfälscht worden wäre.

Neben der durchgeführten Einzelauswertung der Aufgaben inklusive der daraus resultierenden Aktion, fand zudem eine ganzheitliche Betrachtung statt. Zeigten Schüler zum Beispiel ein sich wiederholendes Antwortmuster (alle Kreuzchen links oder alle rechts, symmetrische Kreuzmuster oder ähnliches), so wurde davon ausgegangen, dass diese Schüler nicht teilnehmen wollten. Dies führte ebenfalls zu einem Ausschluss von der Studie.

9.2.2 Anforderungen durch Text als Input

Die Lernwirksamkeit der Repräsentationsform Text war nicht das primäre Forschungsziel. Der hier verwendete Text diente vielmehr als Referenz der Lernwirksamkeit im

Vergleich zu Experten-Concept-Maps. Um eine derartige Vergleichbarkeit zu gewährleisten, werden im Folgenden die relevanten Anforderungen an Schüler durch das Repräsentationsformat Text aufgeführt.

Neben den bereits beschriebenen Leistungen der Lernenden bei der Auseinandersetzung mit Repräsentationen (siehe Lehren und Lernen mit Repräsentationen auf Seite 36 ff.), liegt der Fokus in diesem Fall ausschließlich auf den kognitiven Operation in Bezug auf den eingesetzten Text (ohne Farbleitcode siehe Seite 110, mit Farbleitcode Seite 228).

Aufgrund des hohen Bekanntheitsgrades des Repräsentationsformates Text fällt für Schüler das Erlernen von neuen Routinen durch die Bearbeitung eines Textes sehr einfach aus. Es entspricht allenfalls einer Wiederholung oder Übung von bereits bestehenden Mechanismen (Text lesen und Informationen entnehmen). Entsprechend seiner gekannten Linearität beginnen Lernende mit dem Lesen eines Textes immer automatisch an der richtigen Stelle und wissen genau wo er endet. Dieser an sich triviale Sachverhalt entlastet die Lernenden jedoch beträchtlich, u.a. bei der Erarbeitung neuer Inhalte.

Zusätzlich kann den Lernenden eine Strukturierung von Informationen durch Absätze und Paragraphen innerhalb des Textes zusätzlich entlasten. Ein weiterer Vorteil der Textform gegenüber anderen Repräsentationsformaten wie zum Beispiel Abbildungen ist die Fähigkeit, abstrakte Inhalte zu transportieren. Sobald das der Fall ist, steigen die Anforderungen an den Lernenden. Zwar ist lesen eine kognitive Operation, dennoch geht eine Schwierigkeit mit der Einführung von neuen, unbekanntem (Fach)-Wörtern einher.

Im Vorgriff auf die späteren Lernaufgaben ist es zwingend erforderlich, dass die Lernenden die Bedeutung bestimmter Fachwörter kennen. Ist dies nicht der Fall, so ist eine korrekte Beantwortung der Fragen nicht möglich. Für diese Studie lauteten die neuen Fachwörter: *Sehsinneszellen, Stäbchen, Zapfen, Schaltzellen, Nervenfasern, Netzhaut, lichtempfindlich, Lichtreiz, Sehnerv, Augapfel, blinder Fleck* und *gelber Fleck*.

Entsprechend gängigen lernpsychologischen Forderungen wurde der Text zur besseren Verständlichkeit zusätzlich mit Farben versehen (Mayer 2001, Mayer und Moreno 2003, Schnotz und Bannert 1999, Weidenmann 2006). Die Farben ergeben einen Code, der die Aufmerksamkeit der Lernenden leiten soll. Der Farbleitcode bezieht sich sowohl auf die gleichen Inhalte im Text als auch in den dazugehörigen Abbildungen.

Der Einsatz eines Farbleitcodes kann jedoch auch zu einem Mehr an kognitiver Belastung führen. So müssen Schüler die Kodierung überhaupt erst wahrnehmen, dann müssen sie sie in Kombination mit den Abbildungen bringen, in den Abbildungen wiedererkennen, verstehen, anschließend richtig zuordnen und dann mit der Aufgabe kombinieren. Falls Schüler diese Leistung nicht erbringen, so wirkt die Farbkodierung nicht als Hilfe sondern als Ablenkung.

9.2.3 Anforderungen durch Experten-Concept-Maps als Input

Um eine vergleichbare kognitive Belastung sowohl für den Text als auch für die Experten-Concept-Maps zu erzielen, wurden beide Input-Arten mit der identischen Farbcodierung versehen (ohne Farbleitcode siehe Seite 110, mit Farbleitcode Seite 227). Darüber hinaus galt es, die gleichen lernpsychologischen Anforderungen für die Erstellung der Experten-Concept-Map zu berücksichtigen wie für den Text.

Da beide Arten von bildlichen Darstellungen losgelöst sind, können sie abstrakte Inhalte transportieren. Deshalb wurde bei beiden Input-Arten dieselbe Anzahl an neuen Begriffen verwendet, um die Schwere, die dadurch für die Lernenden entsteht, nicht zu erhöhen.

Im Vergleich zu Text besitzt die Experten-Concept-Map keine vorgegebene Linearität, d.h. keinen eindeutigen Start, kein eindeutiges Ende und keine explizite Leserichtung. Diese Tatsachen stellen für die Lernenden besondere Hürden dar. Nachdem die Lernenden die erste Hürde überwunden haben und eigenständig einen Start gesetzt haben, bedeutet dies nicht automatisch, dass sie die Experten-Concept-Map von Anfang bis Ende lesen können. Hierfür ist die nicht lineare Form verantwortlich, die durch eine Vernetzung der einzelnen Begriffe untereinander mit Richtungswechseln, Sackgassen und Abzweigungen entsteht. Dieser hohe Grad an Vernetzung wirkt sich negativ auf die Übersichtlichkeit aus und kann zudem dafür verantwortlich sein, dass nicht alles von den Lernenden automatisch bearbeitet wird. Weiter bleibt anzumerken, dass der ungewohnte Umgang mit diesem neuen Format dazu führt, dass anfangs mehr Zeit für die Bearbeitung benötigt wird.

Ein weiterer Unterschied der Experten-Concept-Map zum klassischen Text in Bezug auf die zu erbringenden kognitiven Operationen ist, dass sie mit viel weniger klassischem Lesen auskommt. Auf der folgenden Seite werden in Tab. 42 die Gemeinsamkeiten und Unterschiede zwischen Text und Experten-Concept-Map aufgeführt.

Tab. 42: Gegenüberstellung der Anforderungen an Schüler durch Text und Experten-Concept-Map (ECM).

Aussage	Text	ECM
Das Repräsentationsformat ist den Schülern bekannt	Ja	Nein
Die Schüler besitzen bereits Routinen für die Bearbeitung	Ja	Nein
Die Bearbeitung ist linear (Anfang und Ende sind eindeutig)	Ja	Nein
Es gibt eine klare Leserichtung	Ja	Nein
Das Format ist unterteilt und strukturiert (Absätze und Paragraphen)	Ja	Nein
Die Repräsentation kann abstrakte Inhalte vermitteln	Ja	Ja
Die Repräsentation enthält Fach- und Fremdwörter	Ja	Ja
Die Anzahl der Wörter ist...	Hoch	Gering

Dieser Vergleich macht deutlich, dass die beiden Formate zum größten Teil unterschiedliche Anforderungen an Schüler stellen. Unabhängig von Niveau und Inhalt haben beide Formate Vor- und Nachteile. Die Tatsache, dass Experten-Concept-Maps eine geringere Anzahl an Wörtern und geschriebenem Text enthalten, entspricht den Wünschen der Lehrer der Fortbildung und ist somit Teil des Forschungsinteresses als Ansatz zur inneren Differenzierung.

9.2.4 Anforderungen des Multiple-Choice-Tests zur Diagnose

Es wurde ein Testinstrument benötigt, das es ermöglichte, die Menge an transportiertem Fachwissen je Input zu erfassen. Auf den Aufbau und die Struktur wird im Folgenden eingegangen.

Bereits während der Entwicklung lag der Schwerpunkt auf der Diagnose von Schülerleistungen in Bezug auf eine reine Reproduktion von Wissen. Das war nötig, da die Anwendung und der Transfer von Wissen durch die späteren Lernaufgaben im Interview erhoben wurden.

Um zu den bestehenden Fragen des Multiple-Choice-Tests qualifizierte Antworten erstellen zu können, wurden den Schülern die verschiedenen Materialien und der jeweilige Input vorgelegt. Anschließend wurden den Schülern zu jeder Konstellation die bestehenden Multiple-Choice-Fragen als offene Fragen gestellt.

Die hierbei entstandenen Falschantworten wurden umformuliert und als neue Schülergenerierte Distraktoren in den Multiple-Choice-Test integriert. Die auf diese Weise vorgenommene Erhöhung der Antwortmöglichkeiten auf bis zu acht Items, hatte unter anderem zum Ziel, die Ratewahrscheinlichkeit im Multiple-Choice-Test zu reduzieren.

Eine weitere Besonderheit war, dass nicht nur Einfachantworten sondern auch Mehrfachantworten möglich waren und den Schülern zu keinem Zeitpunkt die Anzahl der richtigen Antworten bekannt war.

Bei der Auswertung wurden alle Antworten gezählt. Sowohl die richtig angekreuzten Antworten als auch die richtig nicht angekreuzten Antworten ergaben jeweils einen Punkt. Die erste Aufgabe wurde nicht gewertet, da diese von vorneherein lediglich als Eisbrecher-Aufgabe diente. Die theoretisch aus dem Test resultierende Punktzahl reichte von 0 bis 28 Punkten.

Im Gegensatz zu den halbstandardisierten Dyaden-Interviews wurde bei der Durchführung des Multiple-Choice-Tests auf eine maximale Standardisierung und minimale soziale Interaktion sowohl zwischen dem Versuchsleiter und den Teilnehmenden als auch zwischen den Teilnehmenden untereinander geachtet.

Eine Interpretation der Testergebnisse fiel einfach aus, da die Anzahl der Punkte direkt das Ergebnis widerspiegelte: Je mehr Punkte ein Schüler erzielt hat, desto mehr Informationen konnte er aus dem Input entnehmen (Multiple-Choice-Test siehe Seite 223).

9.2.5 Anforderungen an Schüler durch den Einsatz von Abbildungen und Modellen

Neben den primär eingesetzten Formen des Inputs wurden in dem Interview zwei weitere Repräsentationsformen eingesetzt: Abbildungen und ein Modell vom menschlichen Auge. Wenngleich beide Formen der Darstellung nicht primär von Interesse waren, so erforderte deren Bearbeitung dennoch Leistungen von den Lernenden. Nach dem Input wurden den Lernenden ergänzende Abbildungen zur Erarbeitung gegeben (ohne Farbleitcode siehe Abb. 45, mit Farbleitcode Seite 224).

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 45: Ergänzende Abbildungen zum Input für alle Schüler (ohne Farbleitcode).

Bei den fünf zusammenhängenden Abbildungen in Abb. 45 handelte es sich mehrheitlich um zweidimensionale Darstellungen von Strukturen im menschlichen Auge (entnommen und modifiziert aus Naturwissenschaftliches Arbeiten 3 2005, S. 113). Um den Lernenden zu veranschaulichen, wo sich die Strukturen im Auge befinden, ist in Abbildung 1 (in Abb. 45) ein menschliches Gesicht von der Seite dargestellt. Von dem Auge des dargestellten Gesichtes geht ein Kasten mit Verbindungspfeil in Form

einer Sprechblase weg und enthält die schematische Darstellung eines menschlichen Auges im Längsschnitt. Entsprechend des Farbleitcodes aus dem Text und der Experten-Concept-Map wurden in Abbildung 2 (in Abb. 45) die gleichen Farben für die gleichen Strukturen verwendet. Darüber hinaus ist der Einfall des Lichtes ins Auge und dessen Brechung durch die Augenlinse angedeutet. Die Abbildungen 3, 4 und 5 (in Abb. 43) sind vergrößerte Darstellungen von besonderen Stellen in der Netzhaut: Dem Randbereich der Netzhaut, dem gelben Fleck und dem blinden Fleck. Auch hier finden sich Strukturen, deren Farben denen im Text und in der Experten-Concept-Map entsprechen.

Um ein Verstehen der Abbildungen für Lernende zusätzlich zu erleichtern, sind die Abbildungen sehr vereinfacht dargestellt. Darüber hinaus sind die Abbildungen entsprechend der europäischen Leserichtung von links nach rechts in einem DIN A4 Querformat aufgereiht und mit Sprechblasen als Vergrößerungen versehen. Um den Vergrößerungsschritt von einer Abbildung zur nächsten zu verdeutlichen, sind die Randlinien der Abbildungen pro Vergrößerungsgrad etwas dunkler dargestellt.

Trotz den Modifizierungen zur Verständniserleichterung müssen Lernende dennoch in der Lage sein, den Zusammenhang der Abbildungen und die Vergrößerungsschritte von links nach rechts zu verstehen. Ferner müssen sie die Strukturen getrennt voneinander erkennen bzw. wiedererkennen können. Zusätzlich müssen Lernende auch hier den Farbleitcode wahrnehmen und mit den Inhalten des Inputs in Verbindung bringen. Die Übertragung von sich ergänzenden Inhalten stellt grundsätzlich eine Herausforderung dar und das auch dann, wenn der Input und die Abbildungen zeitlich aufeinander folgen und am Ende direkt nebeneinander vorliegen. Dies verhält sich so, weil der Input und die Abbildungen räumlich voneinander getrennt sind, wodurch das Übertragen von einer Repräsentationsform in die andere erschwert wird (Mayer 2001).

Das Augenmodell

Das Augenmodell, das erst nach der Pilotierung zum Interview hinzugefügt wurde, stellt im Vergleich zu den bereits genannten Repräsentationsformen andere kognitive Anforderungen an die Lernenden.

Als erstes ist der am Modell angebrachte Farbleitcode zu nennen. Dieser entspricht nicht dem bereits in den Abbildungen und im Input verwendeten Farbleitcode. Dieser Umstand konnte nicht behoben werden, da es sich bei dem Modell um kommerzielles Material handelt. Das bedeutete, dass Lernende die Strukturen eigenständig und unabhängig von der Farbe erkennen mussten. Weiter erforderte die hohe Detailtreue des Modells zusätzliche Anstrengungen, da das Mehr an Informationen von den Lernenden auf das Relevante reduziert werden musste. Diese Punkte erforderten eine kognitive

Leistung und führten in der Summe zu erschwerenden Bedingungen für die Lernenden.

Wie bereits zuvor erwähnt, ist die zeitliche Abfolge der Repräsentationsformen zwar gering, allerdings besteht auch hier wieder das Problem der räumlichen Trennung. Erschwerend kommt hinzu, dass am Ende drei Repräsentationen gleichzeitig vor dem Lernenden liegen: Der Input (Text oder Experten-Concept-Map), die Abbildungen und das Augenmodell. Das Modell hat im Vergleich zu den beiden anderen Repräsentationsformen zwei grundlegende Vorteile: Zum einen können Lernende das Modell anfassen, drehen, auseinandernehmen und wortwörtlich begreifen und zum anderen hebt das Modell die zweidimensionalen Inhalte in die räumliche Ebene. Letzteres schien für manche Lernenden eine Belastung darzustellen (umdeuten von 2D in 3D), wohingegen es für andere eine Erleichterung in Bezug auf die Beantwortung der Lernaufgaben war.

9.2.6 Lernaufgaben 1 bis 4 mit kognitiven Anforderungen

Im Folgenden werden die Lernaufgaben aufgeführt, die in Partnerarbeit zu lösen waren. Anschließend werden die kognitiven Operationen, die die Schüler zur Lösung der Aufgaben erbringen mussten, betrachtet.

- **Aufgabe 1:** Wenn du denkst, dass du die Abbildungen verstanden hast, diskutiere mit einem Partner die folgenden Punkte: Betrachtet das Blatt mit dem Auge. Hier seht ihr drei unterschiedliche Abbildungen von der Netzhaut. Vergleicht die Abbildungen 3, 4 und 5 und erklärt, warum im Randbereich der Netzhaut, am gelben Fleck, und am blinden Fleck nicht das Gleiche wahrgenommen werden kann.
- **Aufgabe 2:** Warum siehst du nachts alle Katzen grau?
- **Aufgabe 3:** Es wäre toll, wenn Menschen bei Nacht besser sehen könnten! Was müsste sich deiner Meinung nach am Aufbau der Netzhaut ändern?
- **Aufgabe 4:** Mark hat letztes folgendes erlebt: Es war Nacht, der Mond schien und er fuhr mit seinem Fahrrad einen unbeleuchteten Weg entlang. Leider war das Fahrradlicht kaputt und er hatte Schwierigkeiten den Weg zu erkennen. Anstatt direkt auf den Weg zu blicken, hat Mark zufällig auf den Rand des Weges geschaut und auf einmal konnte er den Weg besser erkennen. Mark konnte sich nicht erklären, warum er bei Dunkelheit den Weg besser erkennen konnte, wenn er nicht direkt darauf geschaut hat. Gib Mark hierfür eine mögliche Erklärung.

Anmerkung: Als gemeinsame kognitive Anforderung für alle folgenden Aufgaben ist Lesen und Verstehen zu nennen.

9.2.6.1 Kognitive Anforderungen von Aufgabe 1

Um diese Aufgabe lösen zu können, mussten die Lernenden zuerst den Input getrennt von den Abbildungen verstanden haben. Erst danach waren sie in der Lage, eine Relation zwischen Input und den Abbildungen herzustellen.

Wie bereits zuvor in diesem Kapitel erwähnt, sind der Input und die Abbildungen nicht redundant, sondern ergänzen sich gegenseitig. Das hat zur Folge, dass die Lernenden bezüglich der Aufgabenstellung zuerst die drei richtigen von den insgesamt fünf Abbildungen identifizieren mussten.

Für ein Lösen der Aufgabe war es weiter erforderlich, Unterschiede im Aufbau der abgebildeten Strukturen zu finden. Hierbei mussten sie erkennen, dass sich im Randbereich der Netzhaut keine Zapfen, sondern nur Stäbchen befinden und dass mehrere Stäbchen mit einer Schaltzelle verbunden sind (Verschaltung 1:viele). Gleichzeitig

sollte ihnen klar werden, dass sich am gelben Fleck nur Zapfen und keine Stäbchen befinden und dass jeder Zapfen mit einer Schatzelle verbunden ist (Verschaltung 1:1), wohingegen sich am blinden Fleck weder Zapfen noch Stäbchen befinden. Aus dem Input haben Lernende entnommen, dass Licht nur mit Sehsinneszellen wahrgenommen werden kann und mussten daraus folgern, dass ohne Sehsinneszellen kein Seheindruck entstehen kann. Ferner sollten sie sich an die Funktionen der Sehsinneszellen erinnern, dass Stäbchen für die Wahrnehmung von hell und dunkel zuständig sind und die Zapfen nur zur Farbwahrnehmung dienen. Schlussendlich mussten Lernende die relevanten Unterschiede in Aufbau und Funktion auf die Wahrnehmung übertragen, abschließend ihre Antwort verbalisieren und im Idealfall noch die neuen Fachwörter verwenden.

Die Musterantwort für Aufgabe 1 lautete: *Im Randbereich der Netzhaut kann nur zwischen hell und dunkel unterschieden werden, da sich hier ausschließlich Stäbchen befinden. Außerdem ist der Seheindruck nicht so scharf wie am gelben Fleck, weil viele Stäbchen ihre Information auf eine Schatzelle weiterleiten. Am gelben Fleck befinden sich nur Zapfen, diese können Farben wahrnehmen. Im Vergleich zum Randbereich der Netzhaut sind die Zapfen 1:1 mit den Schatzellen verschaltet, deshalb kann man hier farbig und sehr scharf sehen. Am blinden Fleck kann man nichts wahrnehmen, da sich hier keine Sehsinneszellen befinden, also keine Stäbchen und keine Zapfen.* Die kognitiven Anforderungen von Aufgabe 1 sind im Folgenden aufgelistet.

Die Schüler müssen

- die Frage lesen und verstehen,
- den Input und die Abbildungen separat voneinander verstehen,
- den Input mit den Abbildungen in Verbindung setzen,
- genau betrachten können,
- die richtige Abbildung finden (welche der fünf?),
- die entsprechenden Strukturen (wieder-) erkennen können,
- verstehen, dass Abbildungen verglichen werden sollen,
- die Abbildungen vergleichen und Unterschiede/Gemeinsamkeiten finden,
- die Unterschiede finden und diese vergleichen können:
 - die Abbildungen betrachten und erkennen, dass die Netzhaut nicht überall gleich aufgebaut ist,
- die Bestandteile der Netzhaut (er)kennen:
 - Randbereich der Netzhaut: keine Zapfen, nur Stäbchen, Verschaltung mit Schatzellen 1:vielen,
 - gelber Fleck: Nur Zapfen, keine Stäbchen, Verschaltung mit Schatzellen 1:1,
 - blinder Fleck: weder Zapfen noch Stäbchen,

- Folgerungen aus den Gemeinsamkeiten/Unterschieden ziehen und in Bezug zur Wahrnehmung setzen,
- den Begriff *Netzhaut* kennen,
- den ungefähren Aufbau des Auges (wieder)erkennen,
- wissen, dass Licht nur mit Sehsinneszellen wahrgenommen werden kann,
- wissen, dass die Abwesenheit von Sehsinneszellen zu keinem Seheindruck führt (blinder Fleck),
- die jeweilige Funktion der Sehsinneszellen wissen:
 - Stäbchen: hell/dunkel-Wahrnehmung,
 - Zapfen: Farb-Wahrnehmung,
- die unterschiedliche Verschaltung der Sehsinneszellen mit den Schaltzellen erkennen,
- das Prinzip der Schaltzellen verstehen:
 - 1: viele = unscharf sehen,
 - 1:1 = scharf sehen,
- die Unterschiede im Aufbau der Netzhaut auf die Wahrnehmung übertragen können,
- ihre Antwort verbalisieren können.

9.2.6.2 Kognitive Anforderungen von Aufgabe 2

Bei der Beantwortung dieser recht kurzen Frage war es von Vorteil, wenn Lernende das Phänomen bereits selbst erlebt hatten.

Sofern dies nicht der Fall war, so mussten sie sich den Sachverhalt zunächst erst vorstellen, was einer zusätzlichen und erheblichen kognitiven Leistung entspricht. Als weiteren Schritt mussten Lernende begreifen, dass mit der in der Aufgabenstellung verwendeten Metapher *Warum siehst du nachts alle Katzen grau?* die Tatsache gemeint ist, dass das menschliche Auge bei Dämmerung Dinge in Grauschattierungen wahrnimmt.

Weiter sollte ihnen klar sein, dass sich die Wahrnehmung nicht allein auf Katzen bezieht. Zur Beantwortung der Frage mussten die Lernenden das Licht als ausschlaggebenden Faktor erkennen und sich gleichzeitig an eine weitere Eigenschaft der Sehsinneszellen aus dem Input erinnern, nämlich dass Stäbchen lichtempfindlicher sind und weniger Licht zum Funktionieren benötigen als die Zapfen. Diese Informationen mussten mit der Funktion hell/dunkel oder Farbwahrnehmung kombiniert und abschließend korrekt verbalisiert werden. Die Musterantwort für Aufgabe 2 lautet: *Da die Zapfen wenig lichtempfindlich sind und nachts nicht viel Licht vorhanden ist, können bei Dunkelheit nur die Stäbchen Licht wahrnehmen. Die Stäbchen können aber*

nur hell/dunkel wahrnehmen, deshalb sehe ich nachts alle Katzen grau. Nachfolgend befindet sich eine Auflistung der speziell durch Aufgabe 2 entstehenden kognitiven Anforderungen.

Die Schüler müssen

- die Frage lesen und verstehen,
- sollten das Phänomen kennen,
- begreifen, dass mit der Aussage die Tatsache gemeint ist, dass man bei Dunkelheit nur in Grauschattierungen sieht (Besonderheit dieser Aufgabe),
- die Lichtverhältnisse bei Dämmerung als ausschlaggebenden Faktor erkennen,
- die Eigenschaften der Sehsinneszellen berücksichtigen:
 - Stäbchen = sehr lichtempfindlich,
 - Zapfen = wenig lichtempfindlich,
- die Funktion der Stäbchen kennen (ermöglichen hell/dunkel Sehen),
- die Funktion der Zapfen kennen (ermöglichen die Wahrnehmung von Farbe),
- die Eigenschaften und Funktionen der Sehsinneszellen in Beziehung zueinander setzen,
- ihre Antwort ausdrücken können.

9.2.6.3 Kognitive Anforderungen von Aufgabe 3

Wie bereits bei den Aufgaben zuvor, mussten Lernende ebenfalls den Aufbau des Auges, die Bestandteile der Netzhaut und deren Funktionen kennen. Jedoch machte die Frage nach einer Verbesserung des Auges ein tiefergehendes Verständnis für die Schwachstellen erforderlich. Ohne solch ein Verständnis waren die Lernenden nicht in der Lage, sinnvolle Verbesserungsvorschläge zu machen.

Die Schwierigkeit der Aufgabe lässt sich wie folgt erklären: Beim Betrachten von Gegenständen fällt das Licht immer auf den gelben Fleck. Dies ist zwar die Stelle des schärfsten Sehens, jedoch befinden sich dort nur Zapfen. Weil aber Zapfen im Vergleich zu Stäbchen mehr Licht benötigen, ist die Wahrnehmung bei Dämmerung eingeschränkt.

Man könnte am gelben Fleck mehr Stäbchen statt Zapfen einbauen. Somit wäre man in der Lage nachts besser sehen zu können, weil die Stäbchen lichtempfindlicher sind als die Zapfen. Weniger Zapfen am gelben Fleck zu haben, würde aber auch bedeuten, dass man am Tag schlechter Farben wahrnehmen könnte.

Im Gegensatz zu den anderen Aufgaben sollte Aufgabe 3 die Lernenden dazu anspornen, ein bestehendes Problem eigenständig anzugehen und unterschiedliche Lösungsmöglichkeiten aufzuzeigen. Denkbare Antworten wären zum Beispiel, wenn Lernende den Malus vom biologischen System Auge als Ganzes erkannt hätten und diesen durch andere Augen wie Eulen ersetzt hätten oder die Augen andersartig umgebaut hätten, damit diese in der Dunkelheit mehr Licht einfangen könnten. Aber auch technische Lösungen waren erdenklich, zum Beispiel durch den Einbau einer Kamera oder eines Restlichtverstärkers. Denkbar wäre auch, den Farbwahrnehmungsbereich der Zapfen zu verbessern oder diese einfach weniger lichtempfindlich zu machen.

Entsprechend der Aufgabenstellung wurden kreative und unorthodoxe Lösungsansätze akzeptiert, sofern diese ausreichend begründet wurden. Sowohl die kognitiven Anforderungen als auch eine Auswahl möglicher Schülerantworten von Aufgabe 3 werden im Folgenden aufgelistet.

Die Schüler müssen

- die Frage lesen und verstehen,
- den Aufbau der Netzhaut kennen:
 - Randbereich der Netzhaut: Keine Zapfen, nur Stäbchen, Verschaltung mit Schaltzellen 1:vielen,
 - Gelber Fleck: Nur Zapfen, keine Stäbchen, Verschaltung mit Schaltzellen 1:1,
- Funktion der Sehsinneszellen wissen:
 - Stäbchen: hell/dunkel-Wahrnehmung,
 - Zapfen: Farb-Wahrnehmung,
- die Eigenschaften der Sehsinneszellen berücksichtigen:
 - Stäbchen sind lichtempfindlich,
 - Zapfen sind wenig lichtempfindlich,
- das Prinzip der Schaltzellen verstehen:
 - 1:vielen = unscharf,
 - 1:1 = scharf,
- begründen können, warum Menschen bei Dunkelheit schlecht sehen können (dies ist bei der Aufgabe nicht explizit gefragt, ist allerdings eine Voraussetzung um die Aufgabe erfolgreich beantworten zu können):
 - Licht fällt hauptsächlich auf den gelben Fleck,
 - Nur am gelben Fleck ist scharfes Sehen möglich,
 - Am gelben Fleck befinden sich allerdings nur Zapfen,
 - Zapfen sind wenig lichtempfindlich, daher sehen Menschen bei Dunkelheit schlecht,

- ihr Wissen anwenden, um sinnvolle Vorschläge zu machen, zum Beispiel
 - den Faktor Licht verändern: Malus vom biologischen System Auge verstehen und diese durch andere Augen (Fisch/Vogel) ersetzen, um mehr Licht einzufangen/ zu spiegeln oder durch technische Lösungen beheben,
 - die Eigenschaften der Sehsinneszellen verändern: Malus der Zapfen erkennen und den Farbwahrnehmungsbereich verbessern oder Malus von Zapfen erkennen und diese mehr lichtempfindlich machen,
 - die Anordnung der Sehsinneszellen verändern: Malus vom Randbereich der Netzhaut erkennen (1:vielen = unscharf) und die Verschaltung verbessern oder Zapfen mit Stäbchen mischen,
- sollten das Probleme kreativ lösen können (das ist das markante an Aufgabe 3).

9.2.6.4 Kognitive Anforderungen von Aufgabe 4

Die letzte Aufgabe bediente sich aller bisherigen fachlichen Inhalte zum Auge. Alleine die Anzahl der zu bedenkenden Faktoren macht diese Aufgabe komplexer und zu dem lag die eigentliche Problemstellung noch als Textaufgabe vor.

Um die Fragestellung mit ihrem bisherigen Wissen zu beantworten, mussten die Lernenden den generellen Aufbau der Netzhaut und besonders die Region des Randbereichs kennen. Ferner mussten sie sich daran erinnern, dass die Stäbchen lichtempfindlicher als die Zapfen sind, keine Farben sehen und nur für die hell-dunkel Wahrnehmung zuständig sind.

Darüber hinaus mussten sie die Situation des Protagonisten in der Geschichte nachvollziehen können und berücksichtigen, dass durch die Tageszeit und das defekte Fahrradlicht schlechte Lichtverhältnisse herrschten und dass dies für die Beantwortung relevant ist. Zusätzlich mussten sie erkennen, dass der Blickwinkel und damit der Lichteinfall zur Beantwortung der Aufgabe mit einbezogen werden musste. Erst dann konnten die Lernenden verstehen, dass das Licht in diesem Fall nicht auf den gelben Fleck, sondern auf den Randbereich der Netzhaut fällt und damit die Aufgabe beantworten.

Diese letzte Aufgabe unterscheidet sich durch die folgenden Merkmale wesentlich von den Aufgaben zuvor: Zum einen die Länge des Aufgabentextes und zum anderen mussten Schüler den Text nicht nur verstehen, sondern auch das Problem erkennen und nachvollziehen können. Zudem stellte alleine die Länge des Aufgabentextes für die Lernenden eine Hürde beim Lösen der Aufgabe dar. Bereits der Blick auf den Text-

körper rief schon während der Pilotierung eindeutig negative Reaktionen bei Schülern hervor.

Analog zu Aufgabe 2 waren auch hier persönliche Erfahrungen von Vorteil, da sich Lernende die Situation sonst zusätzlich vorstellen mussten. Darüber hinaus war es erforderlich, dass sie den geänderten Einfallwinkel des Lichts ins Auge nachvollziehen und diesen in Beziehung zur Wahrnehmung setzen konnten. Für viele Schüler war an dieser Stelle das Augenmodell hilfreich. Die Musterantwort für diese Aufgabe lautet: *Wenn Mark nicht direkt auf den Weg schaut, sondern leicht daneben, dann fällt das Licht des Weges nicht auf den gelben Fleck, sondern auf den Randbereich der Netzhaut. Da sich hier nur Stäbchen befinden und diese weniger Licht als die Zapfen benötigen, kann Mark den Weg bei Dunkelheit besser erkennen, wenn er nicht direkt darauf schaut.* Die kognitiven Anforderungen dieser letzten Aufgabe sind im Folgenden aufgelistet.

Die Schüler müssen

- die Frage lesen und verstehen,
- den Aufbau der Netzhaut kennen:
 - Randbereich der Netzhaut: Keine Zapfen, nur Stäbchen, Verschaltung mit Schaltzellen 1:vielen,
 - Gelber Fleck: Nur Zapfen, keine Stäbchen, Verschaltung mit Schaltzellen 1:1,
- die Eigenschaften der Sehsinneszellen berücksichtigen:
 - Stäbchen = sehr lichtempfindlich,
 - Zapfen = wenig lichtempfindlich,
- die Funktion der Sehsinneszellen wissen:
 - Stäbchen ermöglichen hell/dunkel Sehen,
 - Zapfen ermöglichen Farben-Sehen,
- die Situation nachvollziehen können, um das Phänomen zu verstehen (oder kennen):
 - Nacht = wenig Licht,
 - Fahrradlicht kaputt = wenig Licht,
 - Nicht direkt auf den Weg schauen, sondern leicht daneben,
- die Lichtverhältnisse als ausschlaggebenden Faktor für die Wahrnehmung kennen und den Blickwinkel in ihre Überlegung einbeziehen: Das Licht vom Weg fällt nicht auf den gelben Fleck, sondern auf den Randbereich der Netzhaut (das ist das Besondere an Aufgabe 4),
- ihre Antwort verbalisieren können.

9.2.7 Aufführung der Hilfen zu den Lernaufgaben

Entsprechend den Prinzipien von Aufgaben mit gestuften Hilfen (siehe Seite 87 ff.), sind im Folgenden jeweils die Hilfen für die verwendenden Aufgaben aufgeführt. Zur Erinnerung: Die Hilfen wurden den Schülern im Interview in gefalteter Form ausschließlich dann ausgehändigt, wenn diese von den Schülern explizit gewünscht oder es vom Versuchsleiter als unausweichlich angesehen wurde (siehe Seite 112 ff.). Das geschah nicht in der Reihenfolge der Hilfen sondern entsprechend dem momentanen Problem der Schüler bei den Aufgaben. Für jede der vier Aufgaben wurden für diese Studie fünf Hilfen entwickelt, die den gerade aufgeführten kognitiven Anforderungen der Aufgaben 1 bis 4 entsprechen.

Alle Hilfen sind in zwei aufeinanderfolgende Schritte aufgeteilt. Zuerst sollten die Schüler die Hilfe-Kärtchen einmal auffalten und zuerst die HILFE lesen. Jede HILFE entspricht immer einem Impuls zu einer Aktion. Diese sollten sie ausführen. Wenn die Schüler die Hilfe-Kärtchen ein zweites Mal auffalteten, konnten sie kontrollieren, ob ihre Ausführung richtig war. Bei jeder Aufgabe ist die erste Hilfe die Wiederholung der Aufgabenstellung in eigenen Worten und die letzte Hilfe die Musterlösung der Aufgabe. Im Folgenden sind alle Hilfen für die Aufgaben aufgeführt.

9.2.7.1 Hilfen zu Aufgabe 1

- **HILFE 1:** Erklärt euch gegenseitig die Aufgabe noch einmal mit euren eigenen Worten. Klärt dabei, wie ihr die Aufgabe verstanden habt und was euch noch unklar ist.

ANTWORT 1: Wir sollen die Abbildungen 3, 4 und 5 miteinander vergleichen und Unterschiede suchen. Anhand der Unterschiede sollen wir erklären, warum man im Randbereich der Netzhaut, am gelben Fleck und am blinden Fleck nicht das Gleiche wahrnehmen kann.

- **HILFE 2:** Betrachtet jede der drei Abbildungen noch einmal ganz genau und achtet besonders auf die Farben. Erklärt euch gegenseitig, wie sich die Abbildungen farblich unterscheiden und was die jeweilige Farbe bedeutet (ORANGE, GRÜN, BLAU, GELB).

ANTWORT 2: In allen drei Abbildungen kommt die Farbe ORANGE vor. ORANGE steht für die Ausdehnung der Netzhaut.

- GRÜN steht für die Fortsätze der Schaltzellen, die zusammen den Sehnerv bilden.
- BLAU kommt im Randbereich der Netzhaut vor und steht für die Stäbchen.
- GELB kommt am gelben Fleck vor und steht für die Zapfen.

- **HILFE 3:** Nachdem ihr die unterschiedlichen Strukturen im Auge gefunden habt, solltet ihr wissen, welche Funktionen die Strukturen haben. Lest euch die folgenden Fragen durch und besprecht gemeinsam eure Antwort: (1) Welche Funktionen haben die Stäbchen? (2) Welche Funktionen haben die Zapfen? (3) Was ist die Netzhaut? (4) Wohin führt der Sehnerv? Kleiner Tipp: Sucht nach hilfreichen Informationen im TEXT / in der EXPERTEN-CONCEPT-MAP.

ANTWORT 3: (1) Die Stäbchen können geringe Mengen Licht wahrnehmen und zwischen Hell und Dunkel unterscheiden. (2) Die Zapfen sind nicht so lichtempfindlich wie die Stäbchen. Mit ihrer Hilfe kann man Farben wahrnehmen und unterscheiden. (3) Die Netzhaut befindet sich im hinteren Teil des Auges. Sie ist nicht überall gleich aufgebaut. In der Netzhaut befinden sich Stäbchen, Zapfen und Teile des Sehnervs. (4) Der Sehnerv führt zum Gehirn.

- **HILFE 4:** Ihr habt schon viel über den Aufbau und die verschiedenen Funktionen der Netzhaut herausgefunden. Um die Frage vollständig zu beantworten, müsst ihr auf einen besonderen Unterschied in Abbildung 3 und 4 achten. Sowohl die Stäbchen als auch die Zapfen sind mit weiteren Zellen verbunden. Diese Zellen nennt man Schaltzellen. Was ist der Unterschied bei der Verschaltung zwischen den Sehsinneszellen und den Schaltzellen und was könnte das für das Sehen zur Folge haben?

ANTWORT 4: Im Randbereich der Netzhaut sind mehrere Stäbchen mit einer Schaltzelle verbunden. Der Fortsatz der Schaltzelle zieht in den Sehnerv. Hier verarbeitet eine Schaltzelle die Informationen von bis zu 100 Zapfen und leitet diese zum Sehnerv (1:100). Am gelben Fleck ist jeder Zapfen mit einer Schaltzelle verbunden (1:1). Die direkte Verschaltung einer Sehsinneszelle mit einer Schaltzelle am gelben Fleck führt dazu, dass die Information von einem Zapfen direkt an eine Schaltzelle weitergegeben wird. Durch die direkte Verschaltung am gelben Fleck (1:1) sieht man hier scharf. Deshalb wird der gelbe Fleck auch die Stelle des *schärfsten Sehens* genannt.

- **HILFE 5:** Jetzt habt ihr alles zusammen, um die Aufgabe zu beantworten. Fasst eure gefundenen Unterschiede zusammen und erklärt, warum man im Randbereich der Netzhaut, am gelben Fleck und am blinden Fleck nicht das Gleiche wahrnehmen kann.

ANTWORT 5: Zu Abbildung 3: Im Randbereich der Netzhaut kann nur zwischen hell und dunkel unterschieden werden, da sich hier nur Stäbchen befinden. Außerdem ist der Seheindruck nicht so scharf wie am gelben Fleck, weil viele Stäbchen ihre Information auf eine Schaltzelle weiterleiten. Zu Abbildung 4: Am gelben Fleck befinden sich nur Zapfen, diese können Farben wahrnehmen. Im Vergleich zum Randbereich der Netzhaut sind die Zapfen 1:1 mit den Schaltzellen verschaltet, deshalb kann man hier sehr scharf sehen. Zu

Abbildung 5: Am blinden Fleck kann man nichts wahrnehmen, da sich hier keine Sehsinneszellen befinden (keine Stäbchen und keine Zapfen).

9.2.7.2 Hilfen zu Aufgabe 2

- **HILFE 1:** Dies ist zwar eine kurze Frage, aber auch bei kurzen Fragen kann es Verständnisprobleme geben. Erklärt euch gegenseitig die Aufgabe noch einmal in euren eigenen Worten. Welches Problem (Phänomen) sollt ihr erklären?
ANTWORT 1: Wir sollen erklären, warum wir Menschen in der Dämmerung und bei Nacht Dinge nur noch in Grauschattierungen wahrnehmen können und kaum noch Farben erkennen.
- **HILFE 2:** Diese Hilfe ist einfach. Trotzdem ist es wichtig, den folgenden Unterschied zu kennen: Was ist der Unterschied zwischen Tag und Nacht?
ANTWORT 2: Am Tag scheint die Sonne von morgens bis abends. Nachdem die Sonne untergegangen ist, beginnt die Nacht. Tagsüber gibt es mehr Licht als in der Nacht.
- **HILFE 3:** Mit welchen Zellen im Auge nehmen Menschen ihre Umwelt wahr?
ANTWORT 3: Es gibt zwei Arten von Sehsinneszellen mit denen Menschen sehen können. Das eine sind die Stäbchen und das andere sind die Zapfen.
- **HILFE 4:** Alles klar. Es gibt zwei Arten von Sehsinneszellen. Aber was ist die genaue Funktion der Stäbchen und der Zapfen? Wodurch unterscheiden sie sich? Besprecht eure Antwort.
ANTWORT 4: Die Stäbchen können zwischen hell und dunkel unterscheiden. Im Vergleich zu den Zapfen brauchen die Stäbchen dafür weniger Licht. Die Zapfen können Farben unterscheiden, hierfür benötigen die Zapfen viel Licht.
- **HILFE 5:** Herzlichen Glückwunsch! Ihr habt jetzt alle Informationen zusammen, um die Frage richtig zu beantworten. Tragt noch mal alles was ihr wisst zusammen und formuliert eine Antwort auf die Frage: *Warum sind nachts alle Katzen grau?* Wenn ihr damit fertig seid, trägt einer von euch die Antwort laut vor.
ANTWORT 5: Abends und nachts gibt es weniger Licht. Dies ist nicht genügend Licht, um mit den Zapfen zu sehen. Aber die Lichtmenge ist noch ausreichend, um von den Stäbchen wahrgenommen zu werden. Da die Stäbchen aber nur zwischen Hell und Dunkel unterscheiden können, sehe ich alle Katzen grau.

9.2.7.3 Hilfen zu Aufgabe 3

- **HILFE 1:** Lest euch die Frage genau durch und erklärt euch gegenseitig, was das Problem ist und was ihr machen sollt.

ANTWORT 1: Das Problem ist, dass Menschen bei Nacht schlechter sehen können. Wir sollen uns überlegen was man an dem Aufbau der Netzhaut ändern könnte, damit Menschen in der Dunkelheit besser sehen können.

- **HILFE 2:** Überlegt euch, welche Bestandteile des Auges bei der Lösung des Problems helfen können, so dass man nachts besser sehen kann.

ANTWORT 2: Für das Sehen haben wir zwei Typen von Sehsinneszellen, die Stäbchen und die Zapfen.

- **HILFE 3:** (1) Wohin fällt das Licht im Auge, wenn man etwas direkt anschaut (fixiert)? (2) Welche Zellen befinden sich ausschließlich an dieser Stelle? (3) Was können diese Zellen? Was können diese Zellen nicht?

ANTWORT 3: (1) Das Licht fällt auf den gelben Fleck. (2) Am gelben Fleck befinden sich nur Zapfen. (3) Die Zapfen können Farben wahrnehmen. Die Zapfen brauchen viel Licht um zu funktionieren.

- **HILFE 4:** (1) Welche Zellen brauchen weniger Licht? (2) Was können diese Zellen? Was können diese Zellen nicht?

ANTWORT 4: (1) Die Stäbchen brauchen weniger Licht als die Zapfen. (2) Die Stäbchen können zwischen hell und dunkel unterscheiden. Die Stäbchen können keine Farben wahrnehmen.

- **HILFE 5:** Bevor ihr die Lösung bekommt, sollt ihr noch mal alles zusammenfassen, was ihr bisher gelernt habt und folgende Frage mündlich beantworten: *Was müsste man deiner Meinung nach am Aufbau der Netzhaut ändern, damit Menschen bei Dunkelheit besser sehen können?*

ANTWORT 5: Man könnte am gelben Fleck mehr Stäbchen statt Zapfen einbauen. Da die Stäbchen lichtempfindlicher als die Zapfen sind, könnte man bei Nacht besser sehen. Weniger Zapfen am gelben Fleck zu haben bedeutet aber auch, dass man am Tag schlechter Farben sehen könnte.

9.2.7.4 Hilfen zu Aufgabe 4

- **HILFE 1:** Lest euch die Aufgabe nochmal genau durch. Sagt in euren eigenen Worten, was genau das Problem ist und was ihr erklären sollt.
ANTWORT 1: Mark fährt bei Mondschein Fahrrad und sein Licht ist defekt. Wenn er direkt auf den Weg schaut, erkennt er den Weg nicht so gut, wie wenn er leicht neben den Weg schaut. Mark weiß nicht, warum das so ist und wir sollen ihm das erklären.
- **HILFE 2:** Diese Hilfen sind euch schon bekannt und die Antworten werden euch leicht fallen: (1) Was ist der Unterschied, wenn Mark am helllichten Tag Fahrrad fährt oder bei Mondschein ohne Fahrradlicht? (2) Welche Lichtsinneszellen im Auge von Mark sind eher für die Dunkelheit geeignet?
ANTWORT 2: (1) Am Tag ist es heller als in der Nacht, besonders wenn das Fahrradlicht defekt ist. (2) Da Stäbchen schon geringe Mengen an Licht wahrnehmen können, sind diese geeigneter.
- **HILFE 3:** (1) Wenn Mark direkt auf den Weg schaut (fixiert), wohin fällt dann das Licht auf seiner Netzhaut? (2) Wenn Mark NICHT direkt auf den Weg schaut, wohin fällt dann das Licht auf seiner Netzhaut?

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

ANTWORT 3: (1) Wenn Mark etwas direkt anschaut, dann fällt das Licht auf den gelben Fleck. (2) Wenn er etwas NICHT direkt anschaut, dann fällt das Licht auf eine andere Stelle in der Netzhaut.

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

- **HILFE 4:** Zwischen dem Aufbau der Netzhaut am gelben Fleck und im Randbereich der Netzhaut bestehen Unterschiede, welche sind das?

ANTWORT 4: Am gelben Fleck befinden sich nur Zapfen (Farbwahrnehmung) und im Randbereich der Netzhaut befinden sich nur Stäbchen (Hell-Dunkel-Wahrnehmung).

- **HILFE 5:** Jetzt habt ihr alles Wichtige zusammen und könnt Mark weiterhelfen. Tauscht euch noch mal über das, was ihr gerade gelernt habt, aus und diskutiert eure Antwort. Einer von euch erklärt dann Mark, warum er bei Dunkelheit den Weg besser erkennen kann, wenn er nicht direkt darauf schaut.

ANTWORT 5: Lieber Mark, wenn du etwas direkt anschaust (etwas fixierst), dann fällt das Licht direkt auf den gelben Fleck in deiner Netzhaut. Hier befinden sich nur Zapfen. Diese sind nicht sehr lichtempfindlich und mit ihnen kann man nachts deshalb schlechter sehen. Wenn du etwas nicht direkt anschaust, dann fällt das Licht nicht auf den gelben Fleck, sondern auf den Randbereich in deiner Netzhaut. Hier befinden sich vor allem Stäbchen und diese sind lichtempfindlicher. Mit ihnen kannst du bei Nacht den Weg besser erkennen. Deshalb solltest du nicht direkt auf den Weg schauen, sondern leicht daneben.

9.2.8 Das Erstellen von gestuften Hilfen nach Stäudel (2008)

Zum besseren Verständnis dieser Methode wird in den folgenden Abschnitten das Prinzip hinter Aufgaben mit gestuften Hilfen vorgestellt. Dies beinhaltet sowohl die Erstellung als auch den Einsatz von Aufgaben mit gestuften Hilfen im Unterricht. Abschließend werden die Grenzen dieser Methode aufgeführt, indem gezeigt wird, welche Aufgaben sich für gestufte Hilfen eignen und welche nicht.

Das Erstellen von gestuften Hilfen

Nach Stäudel (Stäudel 2008) zielen die Hilfen zwei Punkte ab: zum einen sollen sie die Schüler inhaltlich unterstützen und zum anderen den Schülern lernstrategische Hilfe geben. Sowohl inhaltliche als auch lernstrategische Hilfen folgen dem gleichen Ablauf. Zuerst wird ein zielgerichteter Impuls gegeben, um zum Beispiel das Vorwissen zu aktivieren. Hierauf folgt eine explizite Ausführung dieses Impulses, sozusagen der richtigen Antwort auf den vorherigen Impuls (konkrete Beispiele hierfür finden sich ab Seite 242).

Eine inhaltliche Hilfe dient dazu, gezielt auf fachliche Aspekte der Aufgabe hinzuweisen, wohingegen lernstrategische Hilfen den Weg zur Lösung anbahnen und das Vorgehen der Schüler strukturieren: *Formuliere die Aufgabe in eigenen Worten!* oder *Suche im Text nach wichtigen Informationen, die du für die Lösung der Aufgabe nutzen kannst* (Stäudel 2008, S. 3).

Je nach Komplexitätsgrad werden pro Aufgabe zwischen 4 und 7 Hilfen entworfen. Die erste Hilfe beinhaltet dabei immer die Aufforderung, die Aufgabenstellung oder das Problem nochmal in eigenen Worten zu formulieren. Hierbei sollen die Schüler sich darüber klar werden, was genau die Frage ist, denn das Verstehen der Aufgabe ist die grundlegende Voraussetzung dafür, das Problem überhaupt bearbeiten zu können. Ferner wird hiermit die sachbezogene Kommunikation der Schüler gefördert.

Die letzte Hilfe entspricht der Musterlösung der Aufgabe. Hiermit können Schüler ihre eigene Lösung vergleichen. Dies dient aber nicht nur der Kontrolle der Richtigkeit der eigenen Ergebnisse, sondern ermöglicht den schwächeren Schülern auf nachvollziehbare Weise zu erkennen, was in dieser Aufgabe von ihnen erwartet wurde und wie sie in Zukunft an ähnliche Aufgaben herangehen könnten (Stäudel 2008, S. 3).

Umgang der Schüler mit den Hilfen

Im regulären Unterricht erhalten die Schüler die Aufgabenstellung und Zugang zu den Hilfen. Damit die Schüler die Aufgabe auch selbst bearbeiten und nicht einfach die Lösung lesen, sollte ihnen der Zugriff auf die Lösung nicht allzu leicht gemacht werden. Das kann zum einen über das Falten der Hilfe-Zettel geschehen (siehe Abb. 46). Zum anderen über die Nachfrage am Pult oder hinter der Tafel. Der Kreativität der Lehrkräfte sind hierbei keine Grenzen gesetzt. Als Gruppengröße für das Arbeiten mit gestuften Hilfen haben sich Zweierteams (Dyaden) bewährt, aber auch Dreierteams oder Einzelarbeit sind möglich.

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

Abb. 46: Faltanweisung der Druckvorlage von Stäudel (2008, Seite 6).

Bei der Einführung von Aufgaben mit gestuften Hilfen muss auf jeden Fall eine Einweisung zur Benutzung der Hilfen gegeben werden. Das beinhaltet den expliziten Hinweis darauf, dass sich die Schüler selbstständig den Zeitpunkt aussuchen, wann sie eine Hilfe in Anspruch nehmen. Ferner sollten Schüler nachdrücklich dazu ermutigt werden, zuerst ernsthaft nach einer eigenen Lösung zu suchen, bevor sie die Hilfen in Anspruch nehmen. Für Hinweise zur Benutzung von gestuften Hilfen gibt Stäudel folgende Anweisung, die als Prototyp einer Einweisung zu sehen ist (Stäudel 2008, S. 5):

- Ihr bekommt jetzt eine Aufgabe, die ihr in Partnerarbeit lösen sollt. Dabei sollt ihr gemeinsam über die Problemstellung sprechen und mögliche Lösungsansätze diskutieren.
- Die Aufgabe enthält alle Informationen, die ihr zum Lösen der Aufgabe braucht. Bearbeitet die Aufgabe so gut, wie es euch möglich ist. Ihr habt für die Bearbeitung der Aufgabe ausreichend Zeit zur Verfügung.
- Wenn ihr mit euren Überlegungen bei der Bearbeitung der Aufgabe nicht weiter kommt, könnt ihr einzelne Hilfen benutzen. Alle Hilfen, die ihr bekommt, beziehen sich auf die Lösung der Aufgabe. Die Hilfen sind also zur Unterstützung gedacht.
- Nehmt jeweils nur eine Hilfe in Anspruch und beginnt dabei mit der Hilfe 1²⁴. Jede Hilfe besteht aus zwei Teilen. Im ersten Teil gibt es entweder eine Frage oder einen Denkanstoß und im zweiten Teil eine Antwort dazu. Ihr sollt zunächst die in der

²⁴ Diese Anweisung entspricht nicht der Verwendung der Hilfen in dieser Studie

Hilfe gestellte Frage bearbeiten bzw. den in der Hilfe gegebenen Denkanstoß gemeinsam diskutieren.

- Überlegt euch, wie ihr die Hilfe zur Lösung der Aufgabe verwenden könnt. Schaut euch erst dann die Antwort zu der Hilfe an. Benutzt die Hilfen in der Reihenfolge ihrer Nummerierung²⁵. Seht euch, auch wenn ihr eine Lösung gefunden habt, die nächsten Hilfen an. Mit den Hilfen könnt ihr nämlich überprüfen, ob eure Lösung richtig ist.

Nicht alle Aufgaben eignen sich für gestufte Hilfen

Um eine Aufgabe mit Hilfen zu versehen, müssen bestimmte Voraussetzungen erfüllt sein. Nach Stäudel sind folgenden Kriterien ausschlaggebend:

- Die Aufgaben sollen ein hohes Niveau haben, sollten aber auch von den leistungstärksten Schülern der Klasse ohne die Hilfen lösbar sein.
- Da die Hilfen auf ein bestimmtes Ziel oder Bearbeitungsweg hin arbeiten, sind prozessorientierte und ergebnisoffene Aufgaben für diese Methode ungeeignet.
- Besser geeignet sind gestufte Hilfen in Verbindung mit praktischem Tun, sofern sie zum Beispiel der Planung oder Ausarbeitung eines Versuchs vorangestellt werden.

Für gestufte Hilfen geeignete Aufgaben sollten komplex, aber eher geschlossen sein. Sie sollten eher der Anwendung dienen als der Neuerarbeitung und sich möglichst nur der Aktivierung und Reorganisation von Vorwissen bedienen (Stäudel 2008, S. 4).

Die in der Studie verwendeten Aufgaben und Hilfen folgen grundsätzlich dem gerade vorgestellten Prinzip, allerdings wurden Modifikationen im Ablauf vorgenommen, um sie dem Format der Studie anzupassen. Die konkreten Aufgabenstellungen und damit verbundenen Anforderungen sowie dazugehörigen Hilfen finden sich ab Seite 112.

²⁵ Die Hilfen wurden in den Interviews nicht in der Reihenfolge ihrer Nummerierung verwendet

9.3 Verwendete Hard- und Software

Alle Concept-Maps in dieser Arbeit und in den dazugehörigen Materialien wurden mit der IHMC CmapsTools (Version 5.06) Software entworfen. Diese kann unter folgender Adresse heruntergeladen werden: <http://cmap.ihmc.us>²⁶. Alle Abbildungen, die nicht einer Concept-Map entsprechen, wurden mit GIMP (Version 2.8.2) entworfen oder verändert. Das Graphikprogramm befindet sich im Downloadbereich von GIMP²⁷. Die Literaturverwaltung und automatische Zitierung fand über Citavi (Version 3.3.0 unter citavi.com), unter Verwendung einer Formatvorlage für WORD 2012 mit Zitationsstil der Humboldt Universität zu Berlin, statt²⁸. Diese Arbeit wurde geschrieben in und formatiert mit Microsoft Office Professional 2010. Die Auswertung der quantitativen Daten erfolgte über IBM SPSS Statistics (Version 19) und dem Freeware Programm *R* (<http://www.r-project.org>). Für die Aufzeichnung der Interviews wurde ein Panasonic Camcorder, Modell HDC-SDX1 verwendet. Die Verschriftlichung der Interviews erfolgte mit der Software MAXQDA (Version 11)²⁹. Alle hier genannten Internetadressen wurden zuletzt geprüft am 18.10.2012.

²⁶ Freeware entsprechend den Lizenzbestimmungen: http://cmap.ihmc.us/download/license_client.php

²⁷ <http://www.gimp.org/downloads/> Freeware entsprechend den Lizenzbestimmungen: LGPLv3+ und GPLv3+

²⁸ http://edoc.hu-berlin.de/e_autoren/word.php?arbeit=Dissertationen&index=index.php&nav=diss

²⁹ <http://www.maxqda.de>

9.4 Ausschreibung der Lehrerfortbildung

Prof. Dr. Ulrike Spörhase
Pädagogische Hochschule Freiburg, Abt. Biologie

Lehrerfortbildung zu Aufgabenentwicklung und Strukturierungsansätzen im Biologieunterricht in der Sek. I

!!! NEUER !!!
ANFANGSTERMIN
16.03.2010 - 14-17Uhr

Aufgaben nehmen eine Schlüsselstellung beim schulischen Lernen ein. Im Biologieunterricht werden fast 70% der Unterrichtszeit mit Aufgaben verbracht und es konnte gezeigt werden, dass diese Zeit besonders lernwirksam ist. Ziel dieser Fortbildung ist die Entwicklung von schulinternen Curricula und Lernaufgaben zur individuellen Förderung der Schüler für den Biologieunterricht. Für die Teilnahme an der Veranstaltung wird ein Zertifikat ausgestellt.

Die Lehrerfortbildung ist Teil eines Forschungsprojektes. Unser Forschungsinteresse ist dabei auf die Vorstellungen der Teilnehmer zur Verbesserung des schulischen Lernens im Biologieunterricht, der Erweiterung ihres Wissens und Könnens, ihren Einstellungen und ihren Fortbildungsintentionen gerichtet.

Zur Konzeption der Lehrerfortbildung

Die Konzeption der Fortbildung geht von den Erfahrungen der Lehrkräfte aus und vermittelt vertieftes fachdidaktisches Wissen und Können zur Entwicklung und zur Strukturierung von biologischen Inhalten. Lehrer und Fachdidaktiker entwickeln gemeinsam schulinterne Lehrpläne mit Modellaufgaben zur inneren Differenzierung für den Biologieunterricht.



Die Inhalte der Lehrerfortbildung umfassen

- Bildungsstandards und Strukturierung von Biologieunterricht
- Bedeutung von Aufgaben im Biologieunterricht
- Basiskonzepte und schulisches Lernen sowie Lernen aus Sicht der Neurobiologie
- Differenzierung (z.B.: nach Vorwissen, Leistungsvermögen, Interessen, Lernertypen)
- Lernaufgaben und Testaufgaben – Niveaunkretisierung, Lernen mit Beispielaufgaben
- Aufgaben und Differenzierung sowie Aufgaben, die das Nachdenken fördern
- Entwicklung von Aufgaben zur individuellen Förderung

Geplanter Ablauf der Lehrerfortbildung

Die Veranstaltung richtet sich an Haupt-, Real- und GymnasiallehrerInnen. Sie beginnt im März 2010 und endet im Herbst 2011 und umfasst ca. 12 Fortbildungstage. Die Fortbildung setzt sich aus halbtäglichen Nachmittagesveranstaltungen und Tagesveranstaltungen zusammen. Die Veranstaltungen werden in den Räumen der PH Freiburg abgehalten. Die Fortbildung wird von Prof. Dr. Ulrike Spörhase und Dr. Martin Otteni geleitet. Eine Fahrtkostenerstattung kann bei den jeweiligen Schülern beantragt werden (Sonderregelung für das Gymnasium).

Geplante Termine für das 1. Halbjahr sind:

Dienstag, den 16.03.2010 halbtags von 14-17 Uhr

Dienstag, den 13.04.2010 ganztägig

Mittwoch, den 16.06.2010 ganztägig

Die Anmeldung erfolgt über Email:
Fragen ist Prof. Dr. Spörhase

Ansprechpartnerin für inhaltliche

9.5 Genehmigung der Studie



Baden-Württemberg
MINISTERIUM FÜR KULTUS, JUGEND UND SPORT

Ministerium für Kultus, Jugend und Sport Baden-Württemberg
Postfach 10 34 42 • 70039 Stuttgart

Pädagogische Hochschule Freiburg
Fakultät für Mathematik, Naturwissenschaften und Technik
Frau Dekanin Prof'in Dr. Spörhase
Kunzenweg 21
79117 Freiburg

Antrag der Pädagogischen Hochschule Freiburg, Fakultät für Mathematik, Naturwissenschaften und Technik, auf Genehmigung einer Studie zur Evaluation von Aufgaben zur individuellen Förderung im Fach Biologie in der Sekundarstufe I in Hauptschulen, Werkrealschulen, Realschulen und Gymnasien

Sehr geehrte Frau Professorin Spörhase,
sehr geehrter Herr Feller,

das Ministerium für Kultus, Jugend und Sport Baden-Württemberg dankt für Ihr Schreiben vom 10. Juni 2011 und genehmigt die vorgelegten aktualisierten Erhebungsinstrumente (Informationsschreiben an die Lehrkräfte und die Erziehungsberechtigten, Einverständniserklärung, Fragebogen).

Das Ministerium wünscht Ihnen einen erfolgreichen Verlauf der Untersuchung.

Mit freundlichen Grüßen

Lambert
Ministerialrat

9.6 Anschreiben an die Lehrkräfte

Sehr geehrte Kollegin, sehr geehrter Kollege.

Die Teilnahme an diesem Projekt (vom Kultusministerium genehmigt) ist freiwillig und es entstehen Ihnen und Ihrer Schule durch die Teilnahme bzw. Nichtteilnahme keinerlei Nachteile. Umso mehr danke ich Ihnen, dass Sie sich an diesem Forschungsprojekt beteiligen. Für den ersten Durchgang in Ihrer Klasse erhalten Sie von mir:

- einen Klassensatz mit Einverständniserklärungen für die Erziehungsberechtigten
- einen Klassensatz mit Fragebögen
- eine Liste mit Fragebogen-Nummern
- einen adressierten und ausreichend frankierten Rückumschlag

Schritt 1: Teilen Sie Ihren Schülern die geplante Teilnahme an einem Forschungsprojekt mit. Verteilen Sie daraufhin die Einverständniserklärungen und erinnern Sie Ihre Schüler, diese baldmöglichst wieder mitzubringen. Sammeln Sie die Einverständniserklärungen zeitnah wieder ein und bewahren Sie diese gut auf. Wenn Sie sich notiert haben, wer alles mitmachen darf, sind Sie bereit für Schritt 2.

Schritt 2: Bevor Sie den Fragebogen austeilen, ist es wichtig, den Schülern folgendes mitzuteilen: *Dieser Fragebogen darf nur von denjenigen ausgefüllt werden, deren Erziehungsberechtigten der Teilnahme zugestimmt haben! Dieser Fragebogen wird nicht bewertet! Bitte vollständige Angaben machen. Es gibt bei diesem Fragebogen kein Richtig oder Falsch, sondern nur eure ehrlich Meinung. Vielen Dank für Eure Mitarbeit.* Das Ausfüllen des Fragebogens dauert circa 20-25 Minuten. Sammeln Sie die Fragebögen wieder ein und bewahren Sie diese gut auf.

Schritt 3: Nachdem Sie die Fragebögen eingesammelt haben, schreiben Sie bitte die Namen der Schüler hinter die entsprechende Fragebogen-Nummer auf der Liste. Schneiden Sie anschließend das Deckblatt von jedem Fragebogen ab und bewahren Sie dieses mit der Kodierungsliste gut auf, das ist sehr wichtig. Packen Sie die ausgefüllten Fragebögen (ohne Deckblatt und Liste!) in den Rückumschlag und schicken Sie mir diesen zurück.

Schritt 4: Nachdem ich die Fragebögen erhalten und ausgewertet habe, melde ich mich bei Ihnen, um den weiteren Ablauf zu besprechen.

Bei Fragen und Anmerkungen können Sie sich jederzeit bei mir melden. Ich danke Ihnen für Ihre Mitarbeit.

9.7 Anschreiben an die Erziehungsberechtigten

Teilnahme an einem Forschungsprojekt

Sehr geehrte Eltern, mein Name ist Wolfgang Feller. Ich bin Doktorand an der Pädagogischen Hochschule Freiburg und erforsche, wie Schüler mit Aufgaben lernen.

Das Ziel meiner Arbeit ist, Aufgaben an die Bedürfnisse von Schülern anzupassen, so dass Schüler trotz unterschiedlicher Voraussetzungen gut lernen.

Nach umfassenden Vorbereitungen möchten wir jetzt die Aufgaben in der Praxis testen. Aus forschungsmethodischen Gründen müssen die Teilnehmenden im Vorfeld einen kurzen Fragebogen ausfüllen. Als nächsten Schritt beschäftigen sich die teilnehmenden Schülerinnen und Schüler mit den Lernaufgaben. Um zu erkennen, welchen Einfluss das Wissen und Können der Schüler hat, müssen die Schüler die Aufgaben bearbeiten und anschließend ihren Lösungsweg erklären. Für die Datenanalyse werden die Schüler hierbei gefilmt.

Ihr Kind darf nur mit Ihrem Einverständnis an dieser Untersuchung teilnehmen, die während einer Schulstunde oder Freistunde stattfindet. Ihr Kind wird dabei nicht bewertet und alle erhobenen Daten werden kodiert und vertraulich behandelt. Die Ergebnisse der Untersuchung werden nur von mir und Mitarbeitern der Pädagogischen Hochschule ausgewertet und sind keiner dritten Person zugänglich. Der Datenschutz ist gewährleistet. Die Teilnahme an der Studie ist freiwillig. Eine Nichtteilnahme ist mit keinerlei Nachteilen verbunden.

Bitte geben Sie Ihrem Kind die rückseitige Einverständniserklärung baldmöglichst wieder mit zurück in die Schule. Für Ihre Zustimmung danke ich Ihnen herzlich. Sie unterstützen damit mein Bemühen, den naturwissenschaftlichen Unterricht weiterzuentwickeln.

Mit freundlichen Grüßen,

Wolfgang Feller

9.8 Augenmodell

In den Interviews wurde ein Augenmodell der Firma 3B Scientific verwendet. Bei dem Modell handelt es sich um ein Auge in 5-facher Vergrößerung, das aus sieben Teilen besteht. Die abnehmbaren Teile beinhalten: Die obere Hälfte der weißen Augenhaut mit Hornhaut und Augenmuskelansätzen, zwei Hälften der Aderhaut mit Iris und Netzhaut, der Linse und dem Glaskörper. Das Auge sitzt auf dem knöchernen Boden der Augenhöhle. Die Abbildungen des Augenmodells sind entnommen von:

http://www.3bscientific.com/Eye-5-times-full-size-7-part-F11,p_1207_65_30_274.html (zuletzt geprüft am 09.09.2013).

Aus urheberrechtlichen Gründen kann die Abbildung in der elektronischen Ausgabe nicht angezeigt werden.

9.9 Weitere Ergebnisse der systematischen Literaturrecherche

- Altersgemischtes Lernen. Eine alltägliche Unterrichtsform an der Jenaplan-Schule Jena (2000). In: Praxis Schule 5 - 10 11 (6), S. 43–48.
- Lernen in jahrgangsgemischten Gruppen (Jg. 3/4/5). [1]. Ein Schulversuch an der Laborschule. Konzeptentwicklung und erste Erfahrungen (2001). Bielefeld: Laborschule.
- Abel, Jürgen; Faust, Gabriele (Hg.) (2010): Wirkt Lehrerbildung? Antworten aus der empirischen Forschung. Münster u.a: Waxmann.
- Brauner, Klaus (2006): Fliegende Superjäger. Experten helfen bei der inneren Differenzierung. In: Schulmagazin 5 bis 10 74 (7-8), S. 45–50.
- Brotman, Jennie S.; Moore Mensah, Felicia; Lesko, Nancy (2011): Urban high school students' learning about HIV/AIDS in different contexts. In: Science Education 95 (1), S. 87–120.
- Caspari, Daniela; Hallet, Wolfgang; Wegner, Anke; Zydatiß, Wolfgang (Hg.) (2007): Bilingualer Unterricht macht Schule. Beiträge aus der Praxisforschung. [Ausgewählte Beiträge vom 21. Kongress für Fremdsprachendidaktik (2005) sowie vom Fremdsprachendidaktischen Kolloquium Berlin-Brandenburg (2006)]. Frankfurt, Main u.a: Lang.
- Doth, Sandra (2013): Was bringen Methoden individualisierten Lernens? Ein Versuch im Biologie/Chemie-Wahlpflichtkurs. In: Pädagogik (Weinheim) 65 (2), S. 11-15.
- Eisenmann, Maria; Grimm, Thomas (Hg.) (2011): Heterogene Klassen - Differenzierung in Schule und Unterricht. 1. Aufl. Baltmannsweiler: Schneider Hohengehren.
- Grotjohann, Norbert; Preisfeld, Gela (2009): Binnendifferenzierung im Biologieunterricht. Seelze: Friedrich (Unterricht Biologie, 347-348/2009).
- Harms, Ute; Bogner, Franz X.; Graf, Dittmar; Gropengießer, Harald; Krüger, Dirk; Mayer, Jürgen et al. (Hg.) (2009): Heterogenität erfassen - individuell fördern im Biologieunterricht. Internationale Tagung der FDdB, Fachsektion Didaktik der Biologie im VBiO, 21. - 25. September 2009. Heterogenität erfassen–individuell fördern im Biologieunterricht. Kiel, 21. – 25. September. Kiel: Leibniz-Institut für die Pädagogik der Naturwissenschaften. Online verfügbar unter <http://www.didaktik-biologie.de/upload/downloads/1255510602.pdf>, zuletzt geprüft am 27.05.2014.
- Heil, Ingeborg; Poot, Peter; Bohrmann, Johannes (2012): Herstellung von Honigwein. Ein fachübergreifendes Rezept. In: Praxis der Naturwissenschaften - Biologie in der Schule 61 (7), S. 38–40.

- Hußmann, Stephan; Selter, Christoph (Hg.) (2013): Diagnose und individuelle Förderung in der MINT-Lehrerbildung. Das Projekt dortMINT. Münster u.a.: Waxmann.
- Jatzwauk, Paul; Rumann, Stefan; Sandmann, Angela (2008): Der Einfluss des Aufgabeneinsatzes im Biologieunterricht auf die Lernleistung der Schüler. Ergebnisse einer Videostudie. In: Zeitschrift für Didaktik der Naturwissenschaften 14, S. 263–283. Online verfügbar unter http://www.ipn.uni-kiel.de/zfdn/pdf/14_014_Jatzwauk.pdf, zuletzt geprüft am 27.05.2014.
- Krüger, Dirk; Meyfarth, Susanne (2009a): Binnen - kurzer Zeit - differenzieren! In: Unterricht Biologie 33 (347/348), S. 2–10.
- Lücken, Markus; Elster, Doris (2009): Überprüfung von Kriterien für den Erfolg von Lerngemeinschaften zur Implementation kompetenzorientierten Unterrichts. In: Ute Harms und u.a (Hg.): Heterogenität erfassen - individuell fördern im Biologieunterricht. Internationale Tagung der FDdB, Fachsektion Didaktik der Biologie im VBiO, 25.-29. September 2009. Kiel: Leibniz-Institut für die Pädagogik der Naturwissenschaften, S. 124–126.
- Menzel, Susanne (2010): Biologische Ressourcen als Lebensgrundlage für alle. Biodiversität als Kontext des Globalen Lernens im Biologieunterricht. In: ZEP : Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik 33 (2), S. 10–15.
- Pädagogisches Landesinstitut Rheinland-Pfalz (2012): Mit Heterogenität umgehen. Differenziert unterrichten in der Sekundarstufe. Fächerübergreifende Anregungen zum Umgang mit Heterogenität im Unterricht. Spezifische Anregungen zum Umgang mit Heterogenität im Fachunterricht. [Sammelband]. Velber: Friedrich.
- Riemeier, Tanja (2005): Kleine Erbse ganz groß. Binnendifferenzierung mithilfe unterschiedlicher Experimentieranleitungen. In: Lernchancen 8 (47), S. 18–25.
- Starzer-Eidenberger, Gudrun (2012): Innere Differenzierung im Biologieunterricht. Konzeption und Umsetzung eines Stationenbetriebs zum Thema Fotosynthese. Salzburg, Univ., Dipl.-Arb., 2012.
- Taxacher, Ingrid (2007): Lernen an Stationen mit innerer Differenzierung im Biologieunterricht. Eine theoriegeleitete Untersuchung anhand von fünf Einzelfallstudien. Salzburg, Univ., Dipl.-Arb., 2007.
- Wellnitz, Nicole; Hartmann, Stefan; Mayer, Jürgen (2009): Evaluation der Bildungsstandards. Entwicklung eines Testinstruments zur modellbasierten Kompetenzmessung. In: Ute Harms, Franz X. Bogner, Dittmar Graf, Harald Gropengießer, Dirk Krüger, Jürgen Mayer et al. (Hg.): Heterogenität erfassen - individuell fördern im Biologieunterricht. Internationale Tagung der FDdB,

Fachsektion Didaktik der Biologie im VBiO, 21. - 25. September 2009.
Heterogenität erfassen– individuell fördern im Biologieunterricht. Kiel, 21. –
25. September. Kiel: Leibniz-Institut für die Pädagogik der
Naturwissenschaften, S. 100–101.

Zentrum für Schulentwicklung (Graz) Österreich / Bundesministerium für Unterricht
und Kulturelle Angelegenheiten (2000): Integration Sekundarstufe 1.
Differenzierter Unterricht in Biologie und Umweltkunde. Klagenfurt: BMUK.

10 EIDESSTATTLICHE ERKLÄRUNG

Vor- und Zuname: Wolfgang Feller

Geburtsdatum: 26.09.1978

Geburtsort: Freiburg im Breisgau

Hiermit versichere ich an Eides statt, dass ich diese Arbeit selbständig angefertigt, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe.

Des Weiteren erkläre ich, dass ich keine weiteren Promotionsversuche unternommen habe.

Wolfgang Feller

Freiburg, den _____