

Molekulare und Zell-Biologie im Schülerlabor
Fachliche Optimierung und Evaluation der Wirksamkeit im
BeLL Bio (Bergisches Lehr-Lern-Labor Biologie)



Dissertation
zur Erlangung des Doktorgrades
der Fakultät der Mathematik und Naturwissenschaften
der
Bergischen Universität Wuppertal

angefertigt am
Lehrstuhl für Zoologie und Biologiedidaktik

vorgelegt von
M.Ed. Karsten Damerau
Wuppertal, im April 2012

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20130312-124320-3

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20130312-124320-3>]

Erstgutachterin & Betreuerin:

Professorin Dr. A. Preisfeld

Zweitgutachter:

Professor Dr. N. Grotjohann

Inhalt

Abstract.....	- 5 -
Zusammenfassung	- 6 -
1 Einleitung	- 7 -
1.1 Fachwissenschaftlicher Hintergrund	- 11 -
1.1.1 Der Genetische Fingerabdruck	- 11 -
1.1.2 Die Photometrische Pyruvat-Bestimmung.....	- 18 -
1.1.3 Einzellige Eukaryoten.....	- 23 -
1.1.4 Biowissenschaftliche Fragestellungen	- 28 -
1.2 Fachdidaktischer Hintergrund	- 29 -
1.2.1 Lernort Schülerlabor.....	- 31 -
1.2.2 Laborvariablen.....	- 38 -
1.2.3 Personenvariablen.....	- 45 -
1.2.4 Experimentieren	- 54 -
1.2.5 Fachdidaktische Fragestellungen.....	- 71 -
2 Material und Methoden	- 77 -
2.1 Fachwissenschaftliche Untersuchungen.....	- 77 -
2.1.1 Molekularbiologische und physiologische Methoden	- 77 -
2.1.2 Photometrische Messungen zum Pyruvatabbau durch die Laktatdehydrogenase ..	- 85 -
2.2 Fachdidaktische Untersuchungen.....	- 87 -
2.2.1 Beschreibung der Stichprobenzusammensetzung.....	- 87 -
2.2.2 Versuchsdesign der Evaluierungen	- 92 -
2.2.3 Die im Rahmen der Evaluierungen eingesetzten Messinstrumente	- 95 -
2.2.4 Überblick über die evaluierten Laborkurse.....	- 118 -
2.2.5 Vorgehensweise bei der Dateneingabe und Auswertung.....	- 135 -
3 Ergebnisse	- 153 -
3.1 Fachwissenschaftliche Ergebnisse	- 153 -
3.1.1 Optimierung des Genetischen Fingerabdrucks	- 153 -
3.1.2 Photometrische Messungen zum Pyruvatabbau durch die Laktatdehydrogenase	- 161 -

3.2	Fachdidaktische Ergebnisse	- 165 -
3.2.1	Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeit (FSWEx)-	165 -
3.2.2	Hauptstudie	- 180 -
3.2.3	Nebenstudie	- 241 -
3.2.4	Lehrerbefragung	- 246 -
4	Diskussion.....	- 252 -
4.1	Fachwissenschaftliche Untersuchungen.....	- 252 -
4.1.1	Der Genetische Fingerabdruck	- 252 -
4.1.2	Photometrische Messungen zum Pyruvatabbau durch die Laktatdehydrogenase	- 258 -
4.2	Fachdidaktisch	- 259 -
4.2.1	Methodendiskussion	- 259 -
4.2.2	Ergebnisdiskussion.....	- 269 -
5	Ausblick	- 287 -
6	Literaturverzeichnis.....	- 288 -
7	Abbildungsverzeichnis.....	- 307 -
Danksagung.....		- 313 -
Anhang		- 314 -
A	Umformulierung der affektiven Items im Posttest für die Kontrollgruppe.....	- 315 -
B	Deskriptive Statistik.....	- 318 -
B.1	Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest – differenziert nach Schülertypen	- 318 -
B.2	Deskriptive Statistik der Laborvariablen im Posttest – differenziert nach Schülertypus	- 320 -
C	Modell affektiver Einflussgrößen im BeLL Bio	- 322 -
C.1	Modell affektiver Einflussgrößen am Lernort BeLL Bio – Direkte Effekte	- 322 -
C.2	Modell affektiver Einflussgrößen am Lernort BeLL Bio – Indirekte Effekte	- 323 -
C.3	Modell affektiver Einflussgrößen am Lernort BeLL-Bio – Totale Effekte	- 324 -
Erklärung		- 325 -

Abkürzungsverzeichnis

ϵ_{λ}	molarer Extinktionskoeffizient ($L \times mol^{-1} \times cm^{-1}$)
ω^2	Omega-Quadrat (Maß der Effektstärke)
η^2	Eta-Quadrat (Maß der Effektstärke)
°C	Grad Celsius
λ	Wellenlänge Lambda
μ	mikro
ΔAbs_{\emptyset}	durchschnittliche Absorptionsänderung
€	Euro
A	Ampere
Abb.	Abbildung
AMOS	Analysis of Moment Structures (Software)
ATP	Adenosintriphosphat
BeLL Bio	Bergisches Lehr-Lernlabor Biologie
bp	Basenpaar
BUW	Bergische Universität Wuppertal
bzw.	beziehungsweise
c	Konzentration (mol/L)
ca.	circa (lat.): ungefähr
CLT	Cognitive Load Theory
CO ₂	Kohlendioxid
CoA	Coenzym A
d	Schichtdicke der Lösung in der Messküvette (cm)
d.h.	das heißt
ddH ₂ O	bidestilliertes Wasser
dest. H ₂ O	destilliertes Wasser

DNA	Deoxyribonucleic acid (engl.): Desoxyribonukleinsäure
DNase	Desoxyribonuklease
dNTP	Desoxynukleotidtriphosphat
EDTA	Etylendiamintetraessigsäure
et al.	et alii/ae/a (lat.): und andere
etc.	et cetera (lat.): und so weiter
<i>F</i>	F-Wert (aus Varianzanalyse)
Fa.	Firma
FADH ₂	Flavin-Adenin-Dinukleotid (reduzierte Form)
FSWEx	Fragebogen zur Erfassung der Selbstwirksamkeit in Bezug auf Experimentiersituationen
<i>g</i>	Erdbeschleunigung
g	Gramm
h	Stunde
H ₂ O	Wasser
HCl	Salzsäure
Ig	Immunglobulin
Ind	Indien
IR	Infrarot
kb	Kilobasenpaare
KMK	Kultusministerkonferenz
L	Liter
LDH	Laktatdehydrogenase
LP	Lehrperson
M	Mittelwert
M	molar (mol/L)
m	Meter
MAC	Macronucleus

Mar	Marokko
ME	Mitteleuropäer
MIC	Micronucleus
min	Minute
MINT	Initialwort für die Fächer Mathematik, Informatik, Naturwissenschaft und Technik
mod.	modifiziert
MSW NRW	Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen
n	nano
N	Stichprobengröße
NAD^+/NADH	oxidiertes/reduziertes Nikotin-Amid-Adenindinukleotid
OE	Osteuropäer
p	Irrtumswahrscheinlichkeit
p	piko
PCR	Polymerase Chain Reaction (engl.): Polymerasekettenreaktion
pH	negativer dekadischer Logarithmus der Protonenkonzentration
PISA	Programme for International Student Assessment
PPP	PowerPoint-Präsentation
r	Korrelationskoeffizient
RFLP	Restriktionsfragmentlängenpolymorphismus
RNase	Ribonuklease
RT	Raumtemperatur
s.	siehe
SD	Standardabweichung
SDS	Natriumdodecylsulfat
STR	short tandem repeat
SuS	Schülerinnen und Schüler

Syr	Syrien
<i>t</i>	t-Wert (aus t-Test)
t	Zeit
Tab.	Tabelle
TAE	Tris-Acetat-EDTA
Taq	<i>Taq</i> -Polymerase (von <i>Thermus aquaticus</i>)
Temp.	Temperatur
TIMSS	Trends in International Mathematics and Science Study
Tris	Tris(hydroxymethyl)-aminomethan
U	Unit (Einheit der Enzymaktivität)
u.a.	unter anderem
UV	Ultravioletter Wellenlängenbereich
V	Volt
vgl.	vergleiche
VIS	visible (engl.): sichtbarer Wellenlängenbereich
VNTR	variable number tandem repeats
W	Watt
w/v	Masse pro Volumen
z. B.	zum Beispiel

Abstract

In Germany, an increasing demand for qualified natural scientists could be observed in the last few years. This demand cannot be covered by the number of graduates in the fields of science and technology. The students' decreasing interest in scientific subjects can already be recorded during their schooldays. Moreover, international comparative studies, such as PISA and TIMSS, report a need for improvement concerning the scientific literacy of German students. Science education that is often rarely practical and that only has low references to everyday life, is deemed as the main reason for this. School laboratories such as the BeLL Bio (Bergisches Lehr-Lernlabor) at the University of Wuppertal seek to counter this trend by offering experiment-based courses that are authentic and have a close proximity to research. Therefore, within the scope of this research study, bioscientific questions, methods and solution strategies were tested and optimized for use in the experiment-based courses of the BeLL Bio. Amongst other things, different human hypervariable DNA-regions were tested for their suitability for use in the school laboratory. These chromosomal gene sequences can, for example, be used to identify suspects in forensics. Moreover, alternatives were examined to substitute ethidium bromide that is often used in molecular biological laboratories, by a substance that is not carcinogenic. Technical procedures of the experiments could be developed further and be established as methods with a close proximity to research for the school laboratory.

The technically optimized courses "Genetic Fingerprinting", "Out of Breath" and "A close look at Protozoa" that evolved from this were then evaluated concerning their affective and cognitive efficiency. In the scope of a main study covering 631 students from the upper secondary level, a significant growth in knowledge, a short rise in students' self concepts concerning their biological capacity and a high authenticity of the laboratory courses could be proofed for all the courses. Moreover, it was possible to compile and evaluate a model of the causal connections between the affective influences at the learning location BeLL Bio. Within the scope of the main study, a theory based and economic questionnaire was developed to analyze the experiment-related perceived self-efficacy (FSWEx). In an ancillary study of 45 students from the upper secondary level, the FSWEx was applied and it could be observed that already a one-time visit to the school laboratory has got the potential to raise the subjective valuation of one's own practical capacity for experimentation. Moreover, it was possible, within the scope of a study involving 45 biology teachers, to identify clear differences in the perceived self-efficacy between teachers feeling confident concerning practical laboratory work and those that do not feel confident. The present study does not only make a significant contribution to the research into effects at the extra-curricular learning location 'students' laboratory', but also to the bioscientific optimization of experiments for courses in school laboratories.

Zusammenfassung

In den letzten Jahren kann in Deutschland ein steigender Bedarf an qualifizierten Naturwissenschaftlern beobachtet werden. Dieser kann durch die Absolventenquoten in den naturwissenschaftlich-technischen Studiengängen nicht gedeckt werden. Bereits in der Schule ist ein abnehmendes Interesse der Schülerinnen und Schüler (SuS) an naturwissenschaftlichen Fächern zu beobachten und internationale Vergleichsstudien wie PISA und TIMSS bescheinigen Schülerinnen und Schülern in Deutschland Optimierungsbedarf in der naturwissenschaftlichen Grundbildung. Als ursächlich dafür wird ein häufig wenig praktischer naturwissenschaftlicher Unterricht mit geringem Alltagsbezug gesehen. Schülerlabore wie das Bergische Lehr-Lernlabor Biologie (BeLL Bio) an der Bergischen Universität Wuppertal bemühen sich, diesem Trend durch forschungsnahe, authentische Experimentierkurse entgegenzuwirken. So wurden im Rahmen der vorliegenden Forschungsarbeit biowissenschaftliche Fragestellungen, Methoden und Lösungsstrategien für den Einsatz in Laborkursen des BeLL Bio optimiert. Unter anderem wurden verschiedene humane, hypervariable DNA-Regionen auf ihre Eignung für den Einsatz im Schülerlabor getestet. Diese chromosomalen DNA-Abschnitte werden beispielsweise zur Identifizierung von Tatverdächtigen in der forensischen Kriminalistik eingesetzt. Weiterhin wurden für die Verwendung im Schülerlabor Alternativen für das in molekularbiologischen Laboren gebräuchliche, cancerogene Ethidiumbromid untersucht. Technische Verfahrensabläufe beim Experimentieren konnten weiterentwickelt und als forschungsnahe Methoden im Schülerlabor etabliert werden. Die daraus entwickelten, fachlich optimierten Laborkurse „Genetischer Fingerabdruck“, „Außer Atem-Sportliche Enzyme“ und „Kleine Einzeller ganz groß“ wurden dann auf ihre affektive und kognitive Wirksamkeit hin evaluiert. Im Rahmen einer Hauptstudie an 631 an Schülerinnen und Schülern der gymnasialen Oberstufe konnte in allen Kursen ein bedeutsamer Lernzuwachs, eine kurzfristige Steigerung des biologischen Fähigkeitsselbstkonzeptes sowie eine hohe Authentizität des Kursangebotes festgestellt werden. Weiterhin gelang es, ein Modell der Kausalzusammenhänge der affektiven Einflüsse am Lernort BeLL Bio zu entwickeln und zu evaluieren. Im Rahmen der Hauptstudie wurde weiterhin ein theoriebasierter, ökonomischer Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung (FSWEx) entwickelt. In einer Nebenstudie an 45 Gymnasiasten wurde unter Einsatz des FSWEx festgestellt, dass bereits ein einmaliger Besuch im Schülerlabor das Potential besitzt, die subjektive Einschätzung der eigenen praktischen Experimentierfähigkeit zu steigern. Darüber hinaus gelang es, in einer Fragebogenerhebung an 45 Biologielehrerinnen und -lehrern, deutliche Unterschiede in der Selbstwirksamkeitserwartung von laborpraktisch sicheren und unsicheren Lehrpersonen zu identifizieren. Die vorliegende Studie leistet damit nicht nur einen

wichtigen Beitrag zur Wirkungsforschung am außerschulischen Lernort Schülerlabor, sondern auch zur biowissenschaftlichen Optimierung von Experimenten in Schülerlaborkursen.

1 Einleitung

„[...] Stimmen aus Handwerk, Mittelstand und Universitäten [reißen nicht ab] [...], die ein mangelndes und weiter sinkendes Qualifikationsniveau der Schulabsolventen beklagen. Ingenieure und Naturwissenschaftler sind Mangelware [...]“ (Klein & Krautz, FAZ 15.03.2012). Gerade in den Bereichen Naturwissenschaften und Technik, die unseren Alltag prägen, wie keine der anderen Wissenschaften, scheint deutlich zu werden, dass in der Ausbildung unseres Nachwuchses qualifizierende und anwendungsorientierte Anteile fehlen. Dabei sind gerade in der Biologie praxisnahe und forschungsrelevante Inhalte unbedingt notwendig. Denn sie bedient mit den Fachanteilen Molekular- und Zellbiologie, Gentechnologie und Physiologie einen Großteil der modernen Disziplinen wie Medizin, Forensik, Biotechnologie und Umweltwissenschaften. Nur müssen diese Inhalte den Schülerinnen und Schülern¹ nahe gebracht werden.

Ein fiktives Szenario: René (17) äußert im schulischen Biologieunterricht: „Und was bringt mir das?“ – Diese oder ähnliche Fragen werden Schüler gelegentlich im Biologieunterricht stellen. Ganz im Sinne der Kommunikationspsychologie können wir dieser Frage vier Botschaften entnehmen (Schulz von Thun 2011), die einige Probleme und gleichzeitig Möglichkeiten des Biologieunterrichts verbildlichen.

Auf der Ebene der Selbstoffenbarung gibt René zu erkennen: Ich erkenne den Nutzen des Unterrichtsinhaltes für mich nicht. Offensichtlich ist es dem Biologieunterricht in dieser Situation nicht gelungen, die Relevanz des Unterrichtsgegenstandes zu vermitteln. Die sachliche Ebene der Frage wirkt vielversprechend, denn sie birgt Potential: René möchte etwas wissen, er möchte dazulernen und etwas über den Nutzen der behandelten Thematik erfahren. Eine Aussage über die Beziehungsebene der Fragestellung zu treffen, ist ohne weitere Hintergrundinformationen schwierig. Sicherlich gilt aber, dass ein gewisses Vertrauensverhältnis zwischen Fragesteller und dem Empfänger, also in der Regel der Lehrkraft, besteht. Von bedeutender Relevanz erscheint die letzte Ebene. Diese beinhaltet den Apell, welcher besagt: Erkläre mir den Nutzen, knüpfe an meine Erfahrungen an.

¹ Im weiteren Verlauf wird aus Gründen der besseren Lesbarkeit auf eine geschlechtliche Differenzierung verzichtet und sowohl für Schülerinnen und Schüler, als auch Lehrerinnen und Lehrer das generische Maskulin verwendet.

Das eben geschilderte Szenario soll die Diskrepanz zwischen bildungspolitischem Anspruch und schulischer Realität verdeutlichen. Die Bildungsstandards im Fach Biologie (KMK 2005) betonen die Bedeutung der Naturwissenschaften für die moderne Gesellschaft. Bereits in der Sekundarstufe I wird die Notwendigkeit gesehen, an Alltagserfahrungen und Vorstellungen der Schüler anzuknüpfen, um diese in Richtung tragfähiger fachlicher Konzepte zu entwickeln. Als günstig wird die Verknüpfung von Lern- und Anwendungsbereichen gesehen, weil sich auf diese Weise fachliches Wissen in sinnstiftenden Zusammenhängen entwickelt (MSW NRW 2008b). Der Biologieunterricht in der gymnasialen Oberstufe soll unter anderem zu genauem praktischen naturwissenschaftlichen Arbeiten anleiten, er soll im Sinne einer wissenschaftspropädeutischen Ausrichtung Einblick in den naturwissenschaftlichen Verständnisweg ermöglichen, an die Studierfähigkeit heranführen und Transferfähigkeit des Wissens auf die Berufswelt erreichen (MSWWF NRW 1999). Diese Anforderungen bemüht sich sicherlich auch jede Lehrkraft nach bestem Wissen und Gewissen zu erfüllen, doch bietet der schulische Rahmen nicht immer die erforderlichen Möglichkeiten. Wie sollen Biologielehrer einen authentischen Einblick in moderne, angewandte Biologie, neue Berufsfelder und Wissenschaft geben, ohne eine authentische Lernumgebung mit forschungsnahen Versuchen bieten zu können? Wie sollen Schüler Methodenkompetenz und Interesse am Fach Biologie entwickeln, ohne praktische, forschungsnahe Experimente selbst durchführen zu können?

Schülerlabore wie das zu den Bergischen Science Labs gehörende Bergische Lehr-Lernlabor der Biologie (BeLL Bio) bieten das Potential, angemessen auf diese Misere zu reagieren. Hier wird Schülern die Möglichkeit geboten, im authentischen Rahmen Versuche durchzuführen, die so im schulischen Unterricht nicht durchführbar wären. Doch ein Laborraum und Experimentieren allein sind noch kein Garant für Motivation und Lernerfolg. Erst die Einbindung der Experimente in ein tragfähiges Unterrichtskonzept verspricht ein erfolgreiches Lernerlebnis (Engeln & Euler 2004). Aufgabe einer experimentell ausgerichteten Biologiedidaktik ist es, diese Konzeption zu gewährleisten. Sie muss Kenntnisse und Methoden aktueller Forschung in didaktisch aufbereiteten Experimenten bereitstellen und deren Lernwirksamkeit überprüfen.

Dieser Aufgabe nimmt sich die folgende Arbeit an. Die lehrplanrelevanten Inhaltsbereiche Molekularbiologie und Zellbiologie werden in experimentellen Laborkursen zu den Themen „Genetischer Fingerabdruck – Dem Täter auf der Spur“, „Außer Atem-Sportliche Enzyme“ und „Kleine Einzeller ganz groß“ umgesetzt. Die Kurse werden im BeLL Bio durchgeführt und auf ihre Wirksamkeit überprüft. Dabei legt die Realisierung der Kurse besonderen Wert auf die forschungsnahe, alltagsrelevante, ungefährliche und ökonomische Anwendung aktueller Labormethoden. Dem letzten Punkt kommt im Kontext der Schülerlaborfinanzierung Bedeutung zu, da die beschriebenen Versuche zum Teil mit erheblichem Aufwand an Verbrauchsmaterial verbunden sind. Insbesondere für den

Laborkurs Genetischer Fingerabdruck machte dies zunächst Untersuchungen zur Optimierung der DNA-Färbung erforderlich, die dann im Schülerlabor Anwendung finden konnten. Die Analyse verschiedener hypervariabler DNA-Loci war notwendig, um reproduzierbare und qualitativ gute Ergebnisse bei der Erstellung des Genetischen Fingerabdrucks zu erzielen. Auf der Suche nach geeigneten Verfahren zur DNA-Präparation für Schülerlaborkurse werden verschiedene Extraktionsmethoden angewandt und modifiziert, um eine schnelle, effektive, ungefährliche und relativ kostengünstige Methode zu etablieren. Im Laborkurs „Außer Atem“ sollte den Schülern aus Gründen der Authentizität der Eindruck vermittelt werden, mit einer (künstlichen) Blutprobe zu arbeiten. Dazu war es notwendig, verschiedene rote Farbstoffe auf eine möglicherweise störende Wirkung der enzymatischen Katalyse photometrisch zu untersuchen.

Wie die notwendigen fachlichen Optimierungen zeigen, ist das volle Potential von Laborkursen noch nicht ausgeschöpft. Doch wie bereits erwähnt, gehört zur Motivation und zum Lernerfolg nicht nur das Experimentieren und ein Laborraum, sondern ebenfalls eine Einbindung in ein didaktisches Konzept (Engeln & Euler 2004). Die tatsächliche Wirksamkeit der Schülerlaborkurse kann aber nur erahnt werden, solange sie nicht überprüft wird. Die gründliche, wissenschaftlich fundierte Evaluation sollte daher integraler Bestandteil solcher außerschulischer Lernorte sein (Brandt et al. 2008). Bisher liegen nur wenige Erkenntnisse über die affektive und kognitive Wirksamkeit der BeLL Bio Laborkurse vor. Generell liegen kaum Erkenntnisse darüber vor, welche Absichten sich mit welchen Experimenten effektiv erreichen lassen (Engeln 2004). Weiterhin sind Studien in Schülerlaboren selten, welche über die Abfrage von allgemeinen Einstellungen von SuS zu Experimenten und Interesse hinausgehen (Brandt 2005). Dem Einfluss von Laborbesuchen auf die kognitive Dimension des Lernens wurde bisher nur wenig (Scharfenberg 2005), dem Einfluss auf die in Experimentalsituationen empfundene Selbstwirksamkeit kaum Beachtung geschenkt. Erstmalig wird in dieser Arbeit eine sehr umfassende Wirkungsanalyse für das noch relativ junge Schülerlabor BeLL Bio vorgenommen, welche sowohl kognitive, als auch verschiedene affektive Auswirkungen eines einmaligen Schülerlaborbesuches berücksichtigt. Gerade weil die Selbstwirksamkeitserwartung, also das aktuelle Empfinden der eigenen Kompetenz, als Prädiktor für Verhalten gilt (Bandura 1982), das gegebenenfalls auch das zukünftige Studiengangs- und Berufswahlverhalten beeinflussen könnte, kommt diesem affektiven Konstrukt in der vorliegenden Arbeit besondere Aufmerksamkeit zu. Im Rahmen der Kursevaluierungen wird daher ein Fragebogen zur Erfassung der Selbstwirksamkeit beim Experimentieren (FSWEx) entwickelt. Dieser erlaubt es, die subjektiv eingeschätzte Experimentierkompetenz von Schülern zu erfassen und hinsichtlich des erzielten Lernzuwachses zu interpretieren. Da die Experimentierfähigkeit der Lehrkräfte als eine mögliche Ursache für die geringe Experimentierhäufigkeit im Unterricht angesehen wird (Scharfenberg 2005), wird der FSWEx zudem im Rahmen einer Befragung von Biologielehrkräften eingesetzt, um deren subjektiv

wahrgenommene Experimentierkompetenz zu messen. Darüber hinaus existieren bisher kaum Untersuchungen, die den Migrationshintergrund (Brandt 2005) oder die berufliche Orientierung der Schüler in die Wirkungsanalyse einbeziehen, weshalb diese Untersuchungsaspekte in der vorliegenden Studie Berücksichtigung finden.

In den weiteren Abschnitten wird zunächst der biowissenschaftliche Hintergrund der molekularen und zellbiologischen Experimente vorgestellt. Anschließend werden die fachdidaktischen Grundlagen der Wirkungsstudie erläutert.

1.1 Fachwissenschaftlicher Hintergrund

Der im Folgenden dargestellte theoretische Hintergrund bezieht sich auf die wissenschaftlichen Experimente zu den Themen „Der Genetische Fingerabdruck“ und „Photometrische Pyruvat-Bestimmung“. Weiterhin werden auch die fachwissenschaftlichen Grundlagen für den Laborkurs „Kleine Einzeller ganz groß“ erläutert. Die fachwissenschaftlichen Untersuchungen sind für eine gelungene Realisierung der genannten Themen im Kontext des Schülerlabors notwendig und Gewinn bringend. Erst diese fachliche Basis eröffnet die Chance, dass Potential des experimentorientierten Unterrichts optimal zu nutzen. Die Vorstellung der Laborkurse selbst erfolgt im Abschnitt „Material und Methode“.

1.1.1 Der Genetische Fingerabdruck

Molekulare Identitätsnachweise, wie sie heute in der Forensik, in der Diagnostik von Krankheiten, beim Vaterschaftstest, aber auch in der Biodiversitätsforschung als DNA-Barcoding oder zum Nachweis gentechnisch veränderter Lebensmittel eingesetzt werden, basieren auf der Feststellung genetischer Unterschiede zwischen verschiedenen Individuen (Kasai et al. 1990; Budowle et al. 1991; Lim et al. 2012; Goldstein & DeSalle 2011). Unter der Berücksichtigung, dass Menschen etwa 99,9% ihres 3,2 Milliarden Basenpaare (bp) umfassenden Genoms teilen, wird die Notwendigkeit deutlich, gezielt interindividuell unterschiedliche Sequenzabschnitte für den Vergleich heranzuziehen. In der Regel werden hierfür als Minisatelliten oder Mikrosatelliten bezeichnete Abschnitte der DNA untersucht (Goodwin et al. 2011). Diese hochrepetitiven Sequenzen machen zusammen etwa 4% des gesamten menschlichen Genoms aus und können von Person zu Person variieren (Abbildung 1; Jasinska & Krzyzosiak 2004). Obwohl viele solcher hochrepetitiven DNA-Abschnitte bekannt sind, ist ihre Bedeutung weiterhin unsicher. Eine begründete Annahme ist, dass sie regulatorische Funktionen bei der Transkription und Expression von Genen übernehmen könnten (Lim et al. 2012).

Minisatelliten – auch VNTRs (*variable number tandem repeats*) genannt – kommen vorwiegend in den subtelomeren Bereichen der Chromosomen vor. Diese sind generell Bereiche, die besonders bei der Modifikation der DNA in Bezug auf epigenetische Regulationen eine Rolle spielen. VNTRs bestehen aus Wiederholungseinheiten zwischen sechs und einhundert Nukleotiden Länge, welche unmittelbar hintereinander bis zu tausendfach wiederholt werden können. Mikrosatelliten, die auch als STRs (*short tandem repeats*) bekannt sind, finden sich überall im menschlichen Genom. Ihre Wiederholungseinheiten sind maximal sechs Nukleotide lang (Goodwin et al. 2011). Eine wichtige Voraussetzung für die Anwendung des genetischen Fingerprintings ist, dass DNA-Konzentrationen im Subnano-Bereich ausreichen müssen, um Individuen damit zu identifizieren (vgl. Kloosterman et al. 1993).

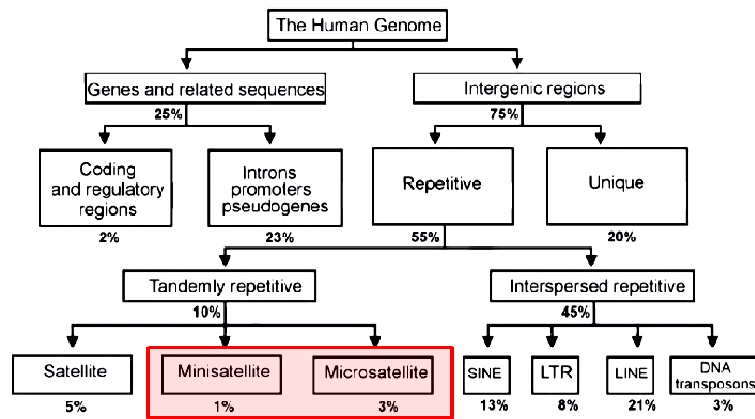


Abbildung 1: Prozentuale Zusammensetzung des menschlichen Genoms; farblich markiert sind Mini- und Mikrosatelliten (verändert nach Jasinska & Krzyzosiak 2004)

Als Beispiel für einen Minisatelliten sei hier der VNTR-Locus D1S80 erwähnt. Er liegt am distalen Ende des p-Arms (kurzer Arm des Metaphasechromosoms) des humanen Chromosoms eins und besteht aus einer Sequenz von sechzehn Nukleotiden Länge. Diese wird zwischen dreizehn (GenBank: AB121699.1) und vierundvierzig (GenBank: AB121730.1) Mal wiederholt, so dass sich zweiunddreißig verschiedene Längenvarianten, Allele genannt, ergeben. Diese sind somit etwa zwischen 200 und 700 Basenpaare lang. Abbildung 2 zeigt exemplarisch die Sequenz einer Wiederholungseinheit des D1S80-Abschnittes. Diese kann im Falle von D1S80 13-44 mal wiederholt werden (Datenquelle: Genebank AB121710.1, Kasai et al. 1990).



Abbildung 2: Nukleotidsequenz eines D1S80 Abschnittes (Datenquelle: Genebank AB121710.1, Kasai et al. 1990).

Während in den Keimzellen des Menschen mit dreiundzwanzig Chromosomen nur ein einfacher Chromosomensatz vorliegt, enthalten alle übrigen menschlichen Zellen einen doppelten Chromosomensatz von sechsundvierzig Chromosomen. Der Partner eines jeden Chromosoms wird als homologes Chromosom bezeichnet. Jeweils eines der homologen Chromosomen wurde vom Vater und eines von der Mutter geerbt (Klug et al. 2007). Somit existiert auch das Chromosom eins sowie der D1S80-Locus in diesen Zellen doppelt und wird nach Mendel vererbt. Dies ist von besonderer Relevanz, da sich die Länge des D1S80-Abschnittes auf diesen beiden homologen Chromosomen in vielen Fällen unterscheidet. Ist die Anzahl der Wiederholungseinheiten und damit die Länge des D1S80-Abschnittes auf beiden homologen Chromosomen identisch, so ist diese Person homozygot für den untersuchten D1S80-Abschnitt (Abb. 3, Genotyp AA oder BB). Variiert die Länge, so ist diese Person heterozygot (Abb.3, Genotyp AB). 86% der Population ist bezogen auf diesen Locus heterozygot, also mischerbig. Eine homozygote Person zeigt im Agarosegel nach erfolgter PCR

(Polymerase-Kettenreaktion zur Amplifizierung der DNA) mit spezifischen Primern für den Locus D1S80 eine Bande, während eine heterozygote Person zwei Banden (1 maternal, 1 paternal) aufzeigt. Die große Variation in den VNTRs kommt möglicherweise durch „normale“ oder fehlerhafte Rekombinationen, Repeats, Punktmutationen, Deletionen oder fehlerhafte DNA-Reparaturen zustande (Ahmad & Seshadri 2007; Das et al. 2002b; Das & Mastana 2003; Duncan et al. 1997; Kloosterman et al. 1993). Aufgrund des polymorphen Charakters eines Locus wie VNTR D1S80 ist es relativ unwahrscheinlich, dass mehrere Menschen dasselbe Muster aufweisen. Dieser Locus besitzt eine Diskriminationsrate von ca. 0,94 (Kloosterman et al. 1993), d.h. die Diversität ist hinreichend hoch, um die Sequenzen zweier Individuen mit hoher Wahrscheinlichkeit voneinander zu unterscheiden. Neben dem D1S80-Locus werden in der vorliegenden Arbeit verschiedene alternative Loci für den Einsatz im Schülerlabor BeLL Bio getestet. Dies sind die Loci D17S30 (VNTR), apoB (VNTR) und Ig – JH (VNTR).

Der VNTR Locus D17S30 befindet sich auf dem Chromosom 17 und umfasst „*tandem repeats*“ einer Core-Sequenz von 70 bp in unterschiedlicher Anzahl in einer Gesamtsequenzlänge von ca. 170-1.000 bp (Wolff et al. 1988; Buscemi et al. 1994). Er ist laut Literatur zum Genetischen Fingerprint ebenso geeignet, weil die Größe der Allele sich ausreichend voneinander unterscheidet, um gut voneinander unterschieden werden zu können (Ivey et al. 1994). Er besitzt z.B. in einer Stichprobe von 202 Personen eine Diskriminationsrate von 0,95 mit einer Ausschließungschance (z.B. bei Vaterschaftsanalysen) von 0,69 (Buscemi et al. 1994). Über 80 % der Kaukasier sind heterozygot in Bezug auf D17S30 (Das & Mastana 2003).

Der VNTR Locus apoB3 sitzt weniger als 100 bp entfernt vom 3'-Ende des Gens für das Apolipoprotein B auf dem humanen Chromosom 2. Das Genprodukt bildet den Proteinanteil für das Lipoprotein des *Low density lipoprotein*-Cholesterol. Lipoproteine erlauben den Transport wasserunlöslicher Lipide im Blut. Die Apolipoproteine befinden sich an der Oberfläche der Lipoproteine und stellen gemeinsam mit Phospholipiden den hydrophilen Bestandteil der Lipoproteine dar. Die apoB3-Core-Sequenz ist 14-16 bp lang und sehr reich an den Nukleotiden Adenin und Thymin. Die Gesamtlänge der PCR-Produkte liegt zwischen ca. 300 und 900 bp. ApoB3 besitzt eine Heterozygotie von ca. 55-78 % und eine Diskriminationsrate zwischen 0,8 und 0,95 (Das et al. 2002a; Das & Mastana 2003; Boerwinkle et al. 1989; Deka et al. 1992a; Frossard et al. 1999).

Der VNTR Locus Ig – J_H sitzt am 5'-Ende der flankierenden Bereiche der schweren Kette des Immunglobulin G-Gens. Immunglobuline (Ig) bestehen aus je 2 schweren (IgH) und zwei leichten Ketten (IgL). Ig – J_H befindet sich auf Chromosom 14, wo sich auch die schweren Ketten flankierende (joining) Region J_H befindet (Decorte et al. 1990; Silva et al. 1987). Der Wert für die Heterozygotie liegt hier zwischen 73 und 82 %. Die „*tandem repeats*“ bestehen aus jeweils 50 Basenpaaren, ihre

Gesamtlänge beläuft sich auf 470 bis 920 bp. Der VNTR-Locus besitzt eine Diskriminationsrate von ca. 0,95 (Silva et al. 1987; Decorte et al. 1990).

Von der isolierten DNA zum genetischen Fingerabdruck

Bevor die beschriebenen DNA-Loci auf ihre Länge hin untersucht werden können, ist es zunächst notwendig, die DNA aus den Zellen zu isolieren. Heute existieren verschiedenste Methoden zur DNA-Isolierung. Neben der Art des zu untersuchenden Probenmaterials hängt die Wahl der Methode auch von ökonomischen Gesichtspunkten, wie der zeitlichen Dringlichkeit und den zur Verfügung stehenden finanziellen Mitteln ab. Weitergehend spielt der unbedenkliche Einsatz im Labor eine Rolle. Trotz der Vielzahl der verwendeten Methoden lassen sich doch Schritte identifizieren, die sich in vielen Methoden wiederfinden lassen. Im ersten Schritt müssen die Zellen aufgeschlossen, also lysiert werden. Anschließend folgt eine Denaturierung der Proteine und im letzten Schritt die Abtrennung der DNA von Proteinen und anderen Zellbestandteilen (Goodwin et al. 2011). Die isolierte DNA kann nun auf Eis gelagert oder direkt für eines der nachfolgend beschriebenen Verfahren eingesetzt werden. Generell finden zwei Verfahren Anwendung, um die Länge der beschriebenen DNA-Abschnitte festzustellen, die Restriktionsfragmentlängenpolymorphismus-Analyse (RFLP-Analyse – engl. *Restriction Fragment Length Polymorphism*) und das PCR-gestützte Verfahren (Teifel-Greding 2000).

Das RFLP-Verfahren

Die RFLP-Analyse ist das historisch ältere Verfahren und beginnt mit einem Verdau der isolierten DNA mit Restriktionsenzymen. Diese schneiden die DNA an spezifischen Restriktionsstellen, so dass in Abhängigkeit vom verwendeten Enzym mehr oder weniger zahlreiche Fragmente entstehen. Deren Länge kann sich von Individuum zu Individuum unterscheiden, da durch Mutationen Restriktionsschnittstellen wegfallen oder hinzukommen können oder weil die Fragmente polymorphe Bereiche wie die oben beschriebenen VNTR's enthalten. Die entstandenen Restriktionsfragmente werden anschließend im Verfahren der Gelelektrophorese entsprechend ihrer Größe aufgetrennt. Zur weiteren Untersuchung wird zunächst ein Southern Blotting durchgeführt. In diesem Verfahren werden die DNA-Fragmente aus dem Gel auf die Oberfläche einer Nylon- oder Nitrozellulosemembran transferiert und mit spezifischen radioaktiven oder nicht radioaktiven Sonden detektiert (Lottspeich & Engels 2006).

Mit dem Aufkommen des PCR-Verfahrens in den 90er-Jahren wurde die RFLP-Analyse im Bereich der Forensik weitgehend durch das schnellere, sehr nachweisempfindliche PCR-Verfahren ersetzt. Da dieses Verfahren auch im Rahmen der experimentellen Untersuchungen Verwendung findet, wird es nachfolgend ausführlich beschrieben.

Die Polymerase-Kettenreaktion (*polymerase chain reaction*, PCR)

Die von Mullis et al. (1986) entwickelte Polymerase-Kettenreaktion (*polymerase chain reaction*, PCR) erlaubt es, aus kleinsten Zell- oder Gewebemengen (selbst *single cell*-PCR ist erfolgreich) solche DNA-Fragmente, deren flankierende Sequenzen bekannt sind, in großen Mengen schnell zu amplifizieren. Mit den PCR-Produkten können dann weitere Untersuchungen folgen, wie z.B. Sequenzanalysen, Sequenzvergleiche oder Nutzung der Sequenz für gentechnische Experimente.

Das Prinzip einer PCR basiert auf der Abfolge vieler Amplifikationszyklen, die jeweils zu einer Verdoppelung der gesuchten DNA-Abschnitte führen. Ein solcher Zyklus besteht aus drei Phasen:

- In Phase eins muss die als Doppelstrang vorliegende DNA bei einer Temperatur von 94 °C denaturiert werden, damit sich die beiden komplementären Stränge trennen.
- Dieses ist eine Voraussetzung für Phase zwei, in der sich kurze Oligonukleotide (Primer, Startermoleküle) anlagern. Für dieses Annealing ist die Kenntnis der flankierenden Regionen erforderlich, da sich deren Sequenz in den verwendeten Oligonukleotiden wieder findet. Die eingesetzten Primer besitzen eine Länge zwischen 18 und 25 bp, sind für jeden Abschnitt spezifisch und dienen in der Reaktion daher als Startermolekül.
- Nach der Anlagerung der Primer an die flankierenden Regionen erfolgt in Phase drei (Elongation), ausgehend von den Primern, die Verlängerung (Extension) des DNA-Stranges von 5'- in 3'-Richtung. Dieser Schritt wird von der *Taq*-Polymerase (von *Thermus aquaticus*, *Taq*), einem temperaturstabilen Enzym, das an die vorliegende Einzelstrang-DNA bindet, katalysiert.

Um sicherzustellen, dass die gesamte DNA von Beginn der Reaktion an in einzelsträngiger Form vorliegt, wird den Zyklen eine 5 minütige Denaturierungsphase vorangestellt. Nach jeder Extension werden die gebildeten Doppelstrang-Fragmente durch Phase 1 im nächsten Zyklus wieder getrennt. Am Ende der PCR wird eine 5 - 10 minütige finale Elongation angehängt, damit alle PCR Produkte vervollständigt werden.

Der Annealingtemperatur kommt für das PCR Ergebnis eine wichtige Bedeutung zu. Ist die Temperatur in der zweiten Phase eines jeden Zyklus zu niedrig gewählt, kann der Primer

unspezifisch binden, was unerwünschte Nebenprodukte nach sich zieht. Wurde sie zu hoch gewählt, kann es sein, dass gar keine Produkte entstehen, da der Primer seinen Schmelzpunkt erreicht und nicht an die DNA-Matritze binden kann. Die maximal zu wählende Annealingtemperatur ist daher abhängig von der Nukleotidsequenz des Primers. Cytosin- und Guanin-reiche Primer vertragen aufgrund von drei Wasserstoffbrückenbindungen gegenüber zwei Bindungen in einer Adenin- und Thymin-reichen Sequenz höhere Temperaturen. Aufgrund der heterogenen VNTR-Sequenzen und vielen unterschiedlichen Individuen muss die optimale Annealingtemperatur für jede Zielsequenz empirisch ermittelt werden (vgl. Erlich 1989, Wittwer et al. 1989, Ferguson et al. 2004, Mülhardt 2009, Goodwin et al. 2011). Für den Einsatz im Schülerlabor gilt es geeignete Primer auszuwählen und PCR-Programme zu optimieren, um eine zuverlässige Amplifikation zu erreichen.

Von der PCR zur Bande im Agarosegel

Nachdem der gewünschte VNTR-Locus in der PCR vervielfältigt wurde, muss die Größe der DNA-Fragmente ermittelt werden. Dies geschieht in der Regel unter Durchführung einer Agarose-Gelelektrophorese. Agarose ist ein Polysaccharid, welches aus roten Meeresalgen gewonnen wird. Beim Aufkochen in Puffer löst es sich und lässt sich in eine Form gießen. Beim Abkühlen erstarrt es zu einem porenhaltigen Gel, welches nach dem Prinzip eines molekularen Siebs arbeitet. Das erstarrte Gel kann anschließend an einem Ende mit den Produkten der PCR-Reaktion beladen werden. Legt man nun ein elektrisches Feld am Gel an, wandert die negativ geladene DNA – also die VNTR-Fragmente unterschiedlicher Länge – durch das Gel in Richtung der Anode. Die kürzeren VNTR-Fragmente werden von dem Gel nicht so stark zurückgehalten und wandern somit schneller durch das Gel, als die längeren Abschnitte. Ein zusätzlich aufgetragener Marker mit DNA-Fragmenten bekannter Größe erlaubt es später, die Größe der in der PCR vervielfältigten VNTR-Fragmente einzuschätzen. Die im Gel zunächst nicht sichtbaren DNA-Fragmente können abschließend durch Färbung mit UV-Farbstoffen, Silberfärbung oder Färbung mit Methylblau als Bandenmuster sichtbar gemacht werden. Alternativ besteht die Möglichkeit, bereits dem noch flüssigen Gel UV-Farbstoffe zuzusetzen (z.B. Ethidiumbromid, Stain G oder GelRed). Diese interkalieren während der Gelelektrophorese in die DNA, so dass die DNA-Banden unmittelbar nach Beendigung der Elektrophorese auf einem UV-Leuchttisch begutachtet werden können (Lottspeich & Engels 2006, Mülhardt 2009, Renneberg 2009).

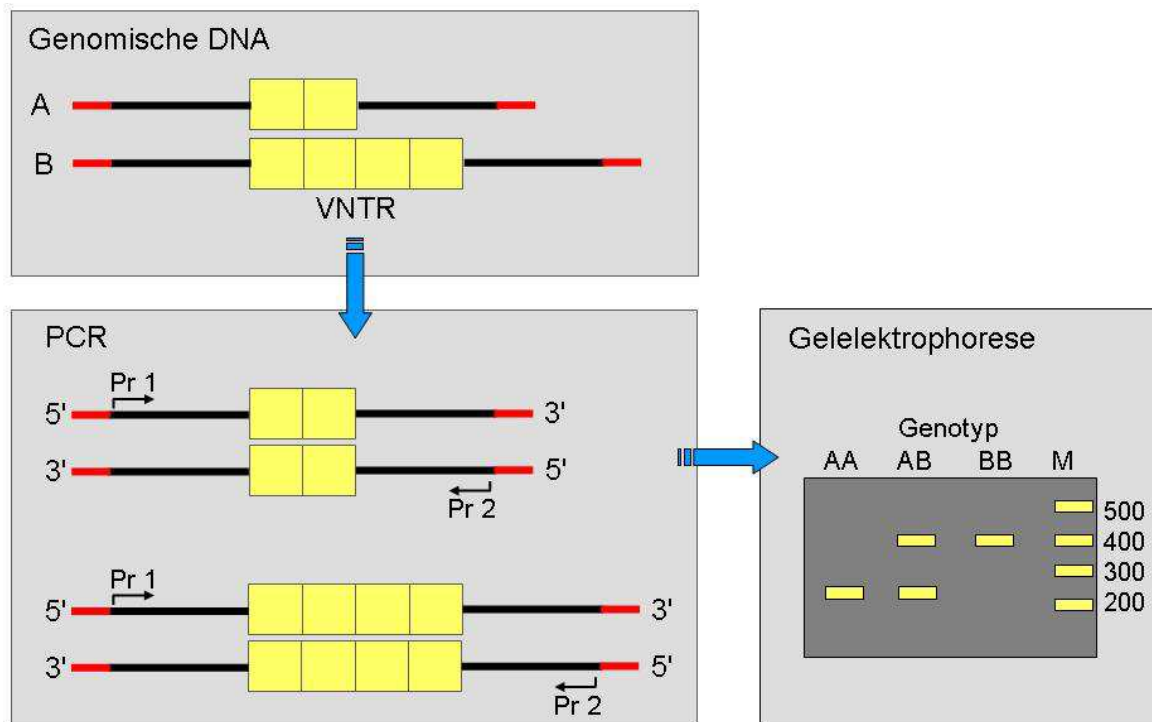


Abbildung 3: Der Weg von der DNA zum Genetischen Fingerabdruck (Pr=Primer; M=DNA-Marker; A=VNTR-Längenvariante A, B=VNTR-Längenvariante B)

Abbildung 3 fasst die bereits beschriebenen Schritte von der genomischen DNA bis zum Gelbild, dem eigentlichen genetischen Fingerabdruck, zusammen. Die genomische DNA wird zunächst isoliert, dann der gewünschte VNTR unter Verwendung spezifischer Primer (Pr1 und Pr2) amplifiziert und die PCR-Produkte in der Gelelektrophorese analysiert. Personen, bei denen sich die Länge des untersuchten VNTR auf beiden homologen Chromosomen unterscheidet, sind heterozygot für den untersuchten Abschnitt (Genotyp AA oder BB). Bei identischer Länge auf beiden homologen Chromosomen ist die Person homozygot für diesen VNTR (Genotyp AB).

1.1.2 Die Photometrische Pyruvat-Bestimmung

Alle lebenden Zellen benötigen für die Aufrechterhaltung von Vitalität und Vermehrung Energie in Form von ATP (Adenosin-5'-Triphosphat). Dazu müssen ihnen permanent Substrate zur Verfügung stehen, die aus dem Blut, der Hämolymphe oder anderen Körperflüssigkeiten aufgenommen oder alternativ aus intrazellulären Speichern abgerufen werden. Der Mensch kann ATP sauerstoffabhängig (aerob) und sauerstoffunabhängig (anaerob) produzieren. Die Entscheidung, ob ein aerober oder ein anaerober Stoffwechselweg eingeschlagen wird, hängt von dem momentanen Bedarf der Zelle und den zellulären Bedingungen ab (Silverthorn 2009).

Beim aeroben Stoffwechsel werden Kohlenhydrate und Fettsäuren unter Sauerstoffverbrauch vollständig zu Wasser (H_2O) und Kohlendioxid (CO_2) abgebaut. Das bei der Glykolyse entstehende Pyruvat (C_3 -Molekül) wird bei Anwesenheit von molekularem Sauerstoff in die mitochondriale Matrix transportiert. Dort wird es mithilfe von Coenzym A decarboxyliert und unter Bildung eines NADH (Nicotinamid-Adenin-Dinukleotid; im Weiteren aus Gründen der besseren Lesbarkeit mit NADH abgekürzt) in das Schlüsselintermediat Acetyl-CoA umgewandelt. Das entsprechende Enzym ist die Pyruvatdehydrogenase. Die C_2 -Acetyleinheit wird nach Abspaltung des Coenzym A in den Citronensäurezyklus eingeschleust, wo es sich mit dem C_4 -Molekül Oxalacetat zum C_6 -Molekül Citrat verbindet (Silverthorn 2009; Abb. 4).

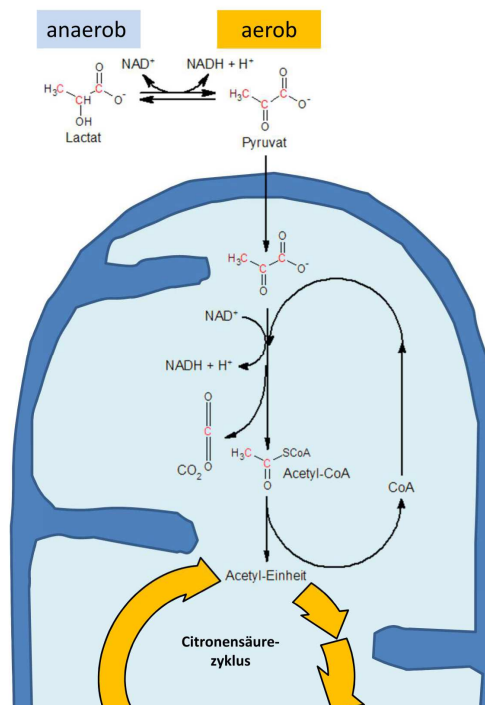


Abbildung 4: Aerober und anaerober Weg des Pyruvats nach der Glykolyse; stark vereinfacht (Eigene Abbildung, in Anlehnung an Silverthorn 2009)

Die im Citronensäurezyklus ablaufenden Schritte lassen aus dem Citrat wieder Oxalacetat entstehen. Die dabei frei werdenden C_1 -Einheiten werden als zwei CO_2 -Moleküle abgegeben. Weiterhin

entstehen pro Umlauf des Zyklus ein ATP, drei NADH und ein FADH₂ (Flavin-Adenin-Dinukleotid). Der Energiegewinn ist beim aeroben Stoffwechsel sehr hoch, da die Glucose vollständig abgebaut werden kann (Abb. 5). Aus einem Glucosemolekül werden bis zu maximal 32 Moleküle ATP gewonnen (Silverthorn 2009).



Abbildung 5: Energiegewinn im aeroben Stoffwechsel (Eigene Abbildung in Orientierung an den Angaben zur ATP-Produktion aus Silverthorn 2009)

Liegt in der Zelle – diese Situation ist besonders häufig in der Muskulatur gegeben – durch intensive oder lang anhaltende Tätigkeit nicht genügend Sauerstoff vor, so wird der anaerobe Stoffwechselweg eingeschlagen. Bei diesen auch als Gärung bezeichneten Abläufen wird Energie aus Glucose ohne Beteiligung von Sauerstoff produziert. Allerdings ist diese Form der Energiegewinnung weniger effektiv, da das Glucose-Molekül nicht vollständig abgebaut wird. Das in der Glykolyse entstandene Pyruvat wird im anaeroben Weg mit Hilfe der Laktatdehydrogenase (LDH) und des Coenzyms NADH zu Laktat (Milchsäure) reduziert, wobei NAD⁺ entsteht (Abb. 6).

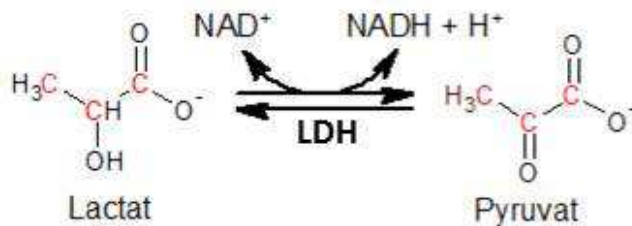


Abbildung 6: Reaktion, welche von der Laktatdehydrogenase (LDH) katalysiert wird (Eigene Abbildung)

Das entstandene NAD⁺ wird in der Glykolyse wieder zu NADH reduziert. Aus einem Molekül Glucose entstehen auf diesem Weg über das Pyruvat 2 Moleküle Laktat, 4 ATP und 2 NADH. Da aber in der Glykolyse auch 2 ATP und 2 NADH verbraucht werden, bleiben als Nettoenergieausbeute nur 2 Moleküle ATP aus einem Molekül Glucose übrig. Die Milchsäure wird zum größten Teil in das Blut abgegeben und teilweise in der Leber und im Herzmuskel wieder zu Pyruvat oxidiert. Für den Umsatz von einem Molekül Pyruvat zu einem Molekül Laktat wird also ein Molekül NADH als Reduktionsäquivalent benötigt.

Die LDH gehört zu den Oxidoreduktasen. Im menschlichen Körper kommen fünf verschiedene Isoenzyme der LDH vor. Isoenzyme sind Enzyme, die dieselbe Reaktion katalysieren, aber sich in ihrer Aminosäuresequenz unterscheiden. Dies führt dazu, dass Isoenzyme unterschiedliche kinetische Parameter aufweisen, was eine Feineinstellung des Metabolismus ermöglicht. Spezifische Isoenzyme sind an die Anforderungen eines bestimmten Gewebes angepasst. Die LDH besteht aus einem Tetramer von Isoenzymuntereinheiten. Die H-Untereinheit wird in der Herzmuskulatur stark

exprimiert und die M-Untereinheit in der Skelettmuskulatur. Verschiedene Kombinationen dieser beiden Untereinheiten zu einem Tetramer sind möglich, so dass das H_4 -, das H_3M -, das H_2M_2 -, das HM_3 - und das M_4 -Isoenzym existieren. Das Isoenzym M_4 kommt schwerpunktmäßig in der Skelettmuskulatur vor und sorgt bereits bei niedrigen Pyruvatkonzentrationen für eine schnelle Reduktion zu Laktat. Im Gegensatz dazu ist das H_4 -Isoenzym auf eine schnelle Oxidation von Laktat zu Pyruvat spezialisiert (Berg et al. 2007, Horn 2009).

In der sportlichen Leistungsdiagnostik wird die Laktatkonzentration im Blut als Anhalt für die sportliche Ausdauerleistungsfähigkeit gemessen (Wahl et al. 2009). Bei einem Belastungsstufentest wird alle 3 min. die Belastungsintensität erhöht und Kapillarblut (z.B. aus Ohrläppchen oder Fingerbeere) entnommen (Saur 2002, Burtcher et al. 2005). Mit diesem Blut lässt sich eine Laktat-Leistungskurve erstellen. Mit zunehmender Belastungsintensität steigt der Laktatgehalt im Blut an (Koch & Raschka 1999). Der Zusammenhang zwischen der Laktatkonzentration und der Belastungsintensität kann Sportwissenschaftlern Hinweise zu Planung des Ausdauertrainings für den Sportler geben. Nach anfangs rein aerober Energiebereitstellung (bis 2 mmol Laktat pro Liter, Schmidt et al. 1997) steigt bei einsetzender Milchsäuregärung die Laktatkonzentration mit der Intensität exponentiell an. Beim gut trainierten Sportler steigt die Laktatkonzentration langsamer an, als bei Untrainierten.

Die Pyruvatkonzentration im Blut liegt gewöhnlich bei unter 0,1 mmol/L. Im gesunden Zustand besteht ein Gleichgewicht aus cytoplasmatischer Bildung von Laktat und Pyruvat sowie deren Verbrauch. Liegen biochemische Defekte auch in vom Pyruvat entfernten Stoffwechselschritten vor, so verändert sich die Konzentration von Pyruvat und Laktat im Blut und im Gewebe. Beispielsweise führt ein Mangel an Pyruvatdehydrogenase zu einer verringerten Einschleusung von Pyruvat in den Citronensäurezyklus und damit zu einer Anreicherung von Pyruvat und Laktat (Smeitink & Wendel 2007). Die Bestimmung des Pyruvatgehaltes im Blut ist somit medizinisch von Relevanz.

Eine Methode zur Bestimmung der Pyruvatkonzentration in Lösung bietet die Photometrie. Photometer werden eingesetzt, um Absorptionen im UV/VIS- bis zum nahen IR-Bereich zu messen (Lottspeich & Engels 2006). Den schematischen Aufbau eines Photometers zeigt Abbildung 7.

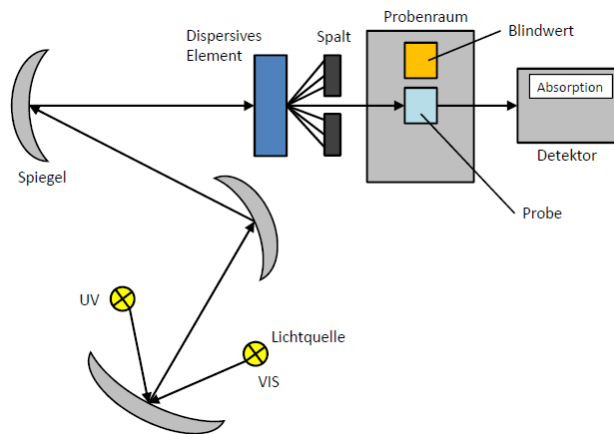


Abbildung 7: Schematischer Aufbau eines Photometers (Eigene Abbildung, in Anlehnung an Richter 2003)

Wesentliche Bestandteile sind ein oder zwei Lichtquellen (für den UV und/oder VIS-Bereich), ein dispersives Element, ein regulierbarer Spalt, eine oder mehrere Halterungen für Proben- und Referenzküvette sowie ein Detektor. Das disperse Element fächert das einfallende Licht in seine Wellenlängen auf. Über einen regulierbaren Spalt wird nur die Wellenlänge selektiert, bei der die photometrische Messung erfolgen soll. Das monochromatische Licht trifft auf die Küvette (Messbehälter) mit der Probenlösung. In der Probe wird das Licht gegebenenfalls absorbiert und das ausfallende Licht von einem Detektor registriert (Richter 2003).

Eine Substanz absorbiert Licht nicht bei allen Wellenlängen gleich gut, wie es beispielsweise vom Photosynthesepigment Chlorophyll a bekannt ist, das seine Absorptionsmaxima bei ca. 420 und 680 nm zeigt (Sadava et al. 2011). Weder Pyruvat noch Laktat absorbieren allerdings im photometrisch erfassbaren Wellenlängenbereich Licht, so dass diese nicht direkt gemessen werden können. Sowohl für NADH, als auch für NAD^+ ist dagegen ein deutliches Absorptionsverhalten im UV-Bereich zu beobachten. NADH zeigt zwei Absorptionsmaxima bei $\lambda=340\text{nm}$ und $\lambda=260\text{nm}$, während sich für NAD^+ lediglich ein Maximum bei $\lambda=260\text{nm}$ findet (Abb. 8).

Das Lambert-Beersche Gesetz erlaubt es, aus der Absorption einer gelösten Substanz bei definierter Wellenlänge auf ihre Konzentration zu schließen.

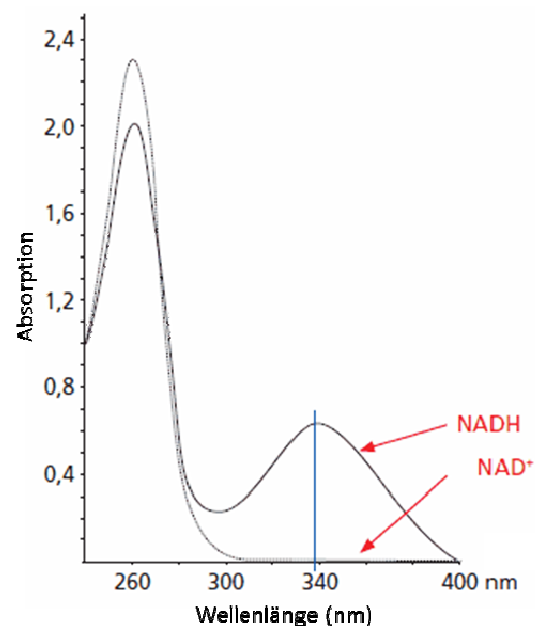


Abbildung 8: Photometrisch aufgenommenes Absorptionsspektrum von NADH und NAD^+ im Wellenlängenbereich von 240-400 nm (verändert nach: <http://www.sigmaaldrich.com/technical-documents/articles/analytix/enzymatic-assay-kits.html>, geprüft am 02.02.2012)

Dabei wird vorausgesetzt, dass keine Lichtstreuung vorliegt, dass die Substanz homogen in Lösung verteilt ist, und dass keine photochemischen Reaktionen in der Lösung erfolgen (Lottspeich & Engels 2006):

$$\Delta Abs = \varepsilon_{\lambda} \times c \times d$$

Lambert-Beersches Gesetz

λ ist das Symbol für die Wellenlänge. Einheit [nm]

ΔAbs ist die Absorption bzw. die Absorptionsänderung bei einer bestimmten Wellenlänge.

ε_{λ} ist der molare Extinktionskoeffizient. Dieser ist eine spezifische Stoffkonstante, die von der Wellenlänge abhängt. Einheit: [L x mol⁻¹ x cm⁻¹]

c ist die Konzentration der absorbierenden Substanz. Einheit: [mol/L]

d ist die Schichtdicke der Lösung in der Messküvette. Einheit: [cm].

Die Absorption des Lösungsmittels wird berücksichtigt, indem eine Referenzprobe (Blindwert) vermessen wird, welche lediglich das Lösungsmittel, nicht aber die zu messende Substanz selbst enthält (Richter 2003). Auf diese Weise ist zwar die Bestimmung der NADH-Konzentration in Lösung, nicht aber die Bestimmung der Pyruvatkonzentration möglich, denn dieses Molekül absorbiert nicht im UV/VIS-Bereich. Wird allerdings ein Überschuss an NADH zu der Pyruvat-Lösung gegeben, so kann das Pyruvat mittels LDH zu Laktat umgesetzt werden. Über Messung der NADH-Absorption vor und nach Umsatz des Pyruvates, anschließende Differenzbildung und Einsetzen der Absorptionsdifferenz in das Lambert-Beersche Gesetz, kann die Ausgangskonzentration an Pyruvat dennoch berechnet werden.

Im Rahmen der Optimierung des Schülerlaborkurses „Außer Atem – Sportliche Enzyme“ wurden Lebensmittelfarbstoffe gesucht, welche es erlauben, die Pyruvatproben rötlich anzufärben, um den authentischen Eindruck einer (künstlichen) Blutprobe zu erwecken. Als geeignet werden Farbstoffe angesehen, die die enzymatische Umsetzung von Pyruvat zu Laktat nicht behindern.

1.1.3 Einzellige Eukaryoten

Die Eukaryoten sind eine der drei Domänen der Organismen, neben den Archaeobacteria und den Bacteria. Entgegen der allgemeinen Annahme, dass vielzellige Organismen, also Pflanzen, Tiere oder Pilze, den Großteil aller eukaryotischen Vertreter ausmachen, weiß man heute, dass das Gegenteil der Fall ist: Die größte Diversität innerhalb der Eukaryoten ist eindeutig den Einzellern oder „Protisten“ zuzuschreiben (Abb. 9).

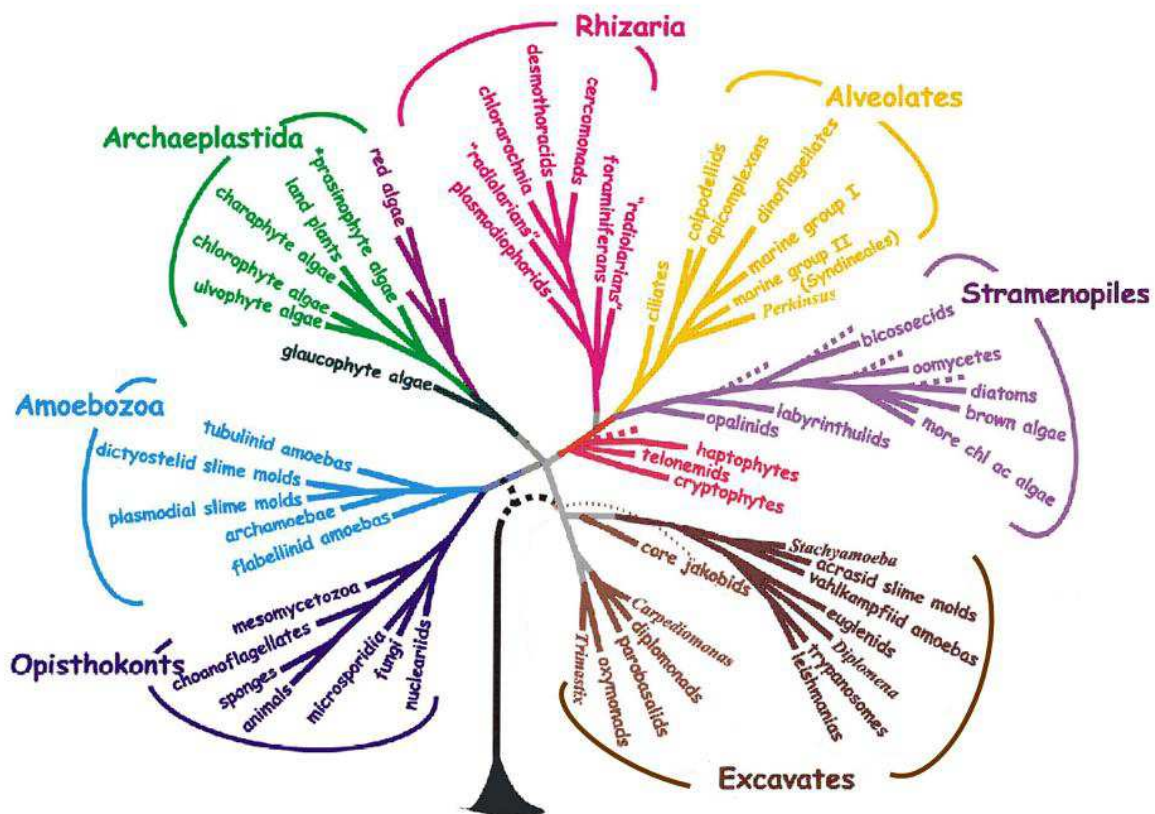


Abbildung 9: Schematischer Stammbaum der Eukaryoten auf der Basis von SSUrDNA-Analysen (verändert nach Baldauf 2008)

Dem Begriff „Protisten“ kommt allerdings keine taxonomische Bedeutung zu, da die als solche zusammengefassten Gruppen polyphyletisch und nicht auf einen gemeinsamen letzten Vorfahren zurückzuführen sind (Abb. 9). Vielmehr kann man die Eukaryoten in sieben distinkte Gruppierungen untergliedern, wobei in allen sieben einzellige Organismen vorkommen (Adl et al. 2005; 2007; Baldauf 2003, Baldauf 2008; Keeling 2010). Allen bekannt sind die Tiere, die sich gemeinsam mit den Pilzen in einem Monophylum befinden, dass als Opisthokonta bezeichnet wird. Die Namensgebung bezieht sich auf eine posterior inserierte, nach vorne den Zellkörper treibende singuläre (opisthokonte) Geißel, wie sie beispielsweise bei den Spermien von Säugetieren vorkommt. Alle dort

eingordneten Organismen weisen in mindestens einem Entwicklungsschritt eine opisthokonte Geißel auf. Weiterhin befinden sich dort die Vorfahren der vielzelligen Tiere, die Choanoflagellaten. Auf einen gemeinsamen letzten Vorfahren mit den Choanoflagellaten und den Tieren greifen auch die Echten Pilze (Fungi) zurück, die allerdings aus traditionellen Gründen von der Botanik behandelt werden. Die Pflanzen bilden zusammen mit den Grünalgen und den Rotalgen die zweite große Gruppe, die auch vielzellige Organismen mit echtem Gewebe umfasst. Dieses als Archaeplastida bezeichnete Taxon bezieht alle Pflanzen und Algen ein, die über eine primäre Endosymbiose (genauer: Endocytobiose) in den Besitz von Plastiden gekommen und nun zur Photosynthese befähigt sind (Keeling 2010; Weber & Osteryoung 2010). Eine weitere Gruppe, in der sich auch mehrzellige Algen befinden, sind die Stramenopila, zu denen die großen als Meerestange bezeichneten Braunalgen, aber auch die einzelligen Kieselalgen gehören. Auch hier liegt als gemeinsames morphologisches Merkmal eine Begeißelung in irgendeinem Entwicklungsstadium vor, nur ist es hier eine heterokonte Begeißelung. Die apikale Geißel ist mit kleinen Härchen, den Mastigonemata, besetzt, die andere ist kürzer und als Schleppgeißel nach hinten gerichtet (Adl et al. 2005). Viele Mitglieder dieser Gruppe haben die Fähigkeit zur Photosynthese sekundär wieder verloren. Alle anderen vier eukaryotischen Großgruppen bestehen ausschließlich aus einzelligen Organismen. Die oftmals mit den Stramenopila zu den Chromalveolata zusammengefassten Alveolaten enthalten die Dinoflagellaten, von denen manche für die bekannten Roten Tiden oder Fischvergiftungen sorgen können, und die durch das Pantoffeltierchen vielen bekannten Wimpertierchen (Ciliophora). Weitere Großgruppen aus Einzellern sind neben den Rhizaria und Amoebozoa noch die Excavata. Zu den Excavata gehören auch die in der Arbeitsgruppe Preisfeld an der Bergischen Universität Wuppertal beforschten Euglenozoa mit sowohl zur Photosynthese befähigten als auch phagotroph lebenden Einzellern (Busse & Preisfeld 2002; 2003; Cavalier-Smith 2002; Adl et al. 2005; 2007).

Nach dieser allgemeinen Einführung in die Großgruppen der Eukaryoten wird im Folgenden auf die Einzeller eingegangen, die im Schülerlaborkurs „Kleine Einzeller ganz groß“ eingesetzt wurden. Dieses sind *Tetrahymena thermophila* (Ciliophora, Wimpertierchen) und *Euglena gracilis* (Euglenida, Augentierchen). Obwohl die deutschen Namen beider Einzeller auf „Tierchen“ enden, haben sie keine nähere Verwandtschaft mit den Tieren. Sie sind weder den Tieren, noch den Pflanzen oder Pilzen zuzuordnen, sondern befinden sich innerhalb der Alveolata sowie der Excavata (Abb. 9).

***Tetrahymena thermophila* (Ciliophora)**

T. thermophila ist ein farbloser Ciliat von 30-60 µm Größe (Abb. 10) und ubiquitärer Verbreitung. Seine Oberfläche ist mit zahlreichen Wimpern (Cilien) bedeckt, die sowohl der Fortbewegung als auch dem Nahrungserwerb durch Heranstrudeln von Nahrungspartikeln dienen.

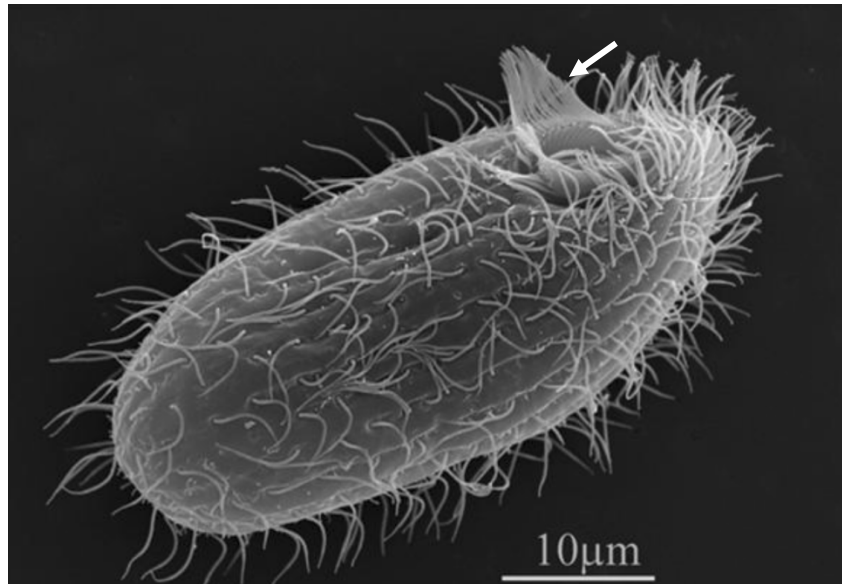


Abbildung 10: Raster-Elektronenmikroskopische Aufnahme von *Tetrahymena thermophila*. Pfeil: Mundfeld mit undulierender Membran (<http://www.miami.muohio.edu/news/media/1588.jpg>, geprüft am 23.03.12)

Tetrahymena ernährt sich über Phagozytose, d.h. Nahrungspartikel werden im Mundfeld in Phagosomen aufgenommen und für die weitere Verdauung mit Acidiosomen und Lysosomen verschmolzen, in denen sich Enzyme zur Verdauung der aufgenommenen Nahrung befinden. Nach der Aufnahme der Nährstoffe ins Cytoplasma werden unverdauliche Stoffe über den Cytoproct (Zellafter) ausgeschieden. Die Osmoregulation erfolgt über pulsierende (kontraktile) Vakuolen.

Ein weiteres besonderes Merkmal der Wimpertierchen ist der Kerndualismus, auch Kerndimorphismus genannt. Bei den meisten Ciliaten kommt ein Kerndualismus vor mit einem Micro- und einem Macronucleus. Der polyploide Macronucleus (MAC) enthält zahlreiche Kopien der DNA (ca. 45) und stellt als höchst effektiver Synthese- und Expressionsapparat den somatischen, transkriptionsaktiven Kernanteil dar. Der Micronucleus (MIC) ist diploid und enthält alle genetischen Informationen für die Keimbahn in seinen fünf Chromosomen. Er bleibt während der Wachstumsphase von *Tetrahymena* transkriptionsinaktiv (Karrer 1999, Collins & Gorovsky 2005).

Die als Konjugation bezeichnete sexuelle Fortpflanzung von Ciliaten (hier an *Tetrahymena* erläutert) ist ein ungewöhnlicher und komplexer Prozess, in dem sich zunächst zwei Zellen unterschiedlichen Paarungstyps nach Auslösung eines Hungersignals aneinander lagern (Abb. 11). Im Labor kann man

diese Stadien in nahrungsdefizienten Kulturen nach ca. 30 min. beobachten. Der dann stattfindende Austausch des genetischen Materials beginnt mit den beiden meiotischen Teilungen des MIC zu vier haploiden MICs (1). Drei der haploiden MICs migrieren zum posterioren Ende und degenerieren (2). Der vierte MIC macht eine weitere, postmeiotische Teilung durch. Dieses führt also in jeder Zelle zu zwei haploiden Tochter-MICs bei gleichzeitiger Ausbildung einer Konjugationsbrücke zwischen den beiden Zellen. Je einer der beiden MICs verbleibt, während der andere als Wanderkern über die Plasmabrücke in die andere Zelle wandert (3). Die beiden Tochter-MICs unterschiedlicher genetischer Ausprägung verschmelzen jeweils in den Zellen, wodurch pro Zelle ein neuer diploider MIC entsteht (4). Dieser teilt sich wieder zweimal mitotisch in vier diploide MICs (5). Zwei der MICs wandern zum posterioren Ende, einer löst sich auf, während der andere zum neuen funktionellen MIC mit genetisch rekombiniertem Material wird. Die beiden übrig bleibenden MICs werden zur Makronucleus-Anlage (6). Die Zellen teilen sich nun und der alte MAC degeneriert. Bei der nächsten Teilung entstehen vier haploide identische Tochterzellen mit je einem MIC und dem aus der Macronucleus-Anlage entstandenen neuen MAC (7), womit die sexuelle Reproduktion abgeschlossen ist. Neben dieser kommen auch asexuelle Teilungen vor (Karrer 1999, Collins & Gorovsky 2005).

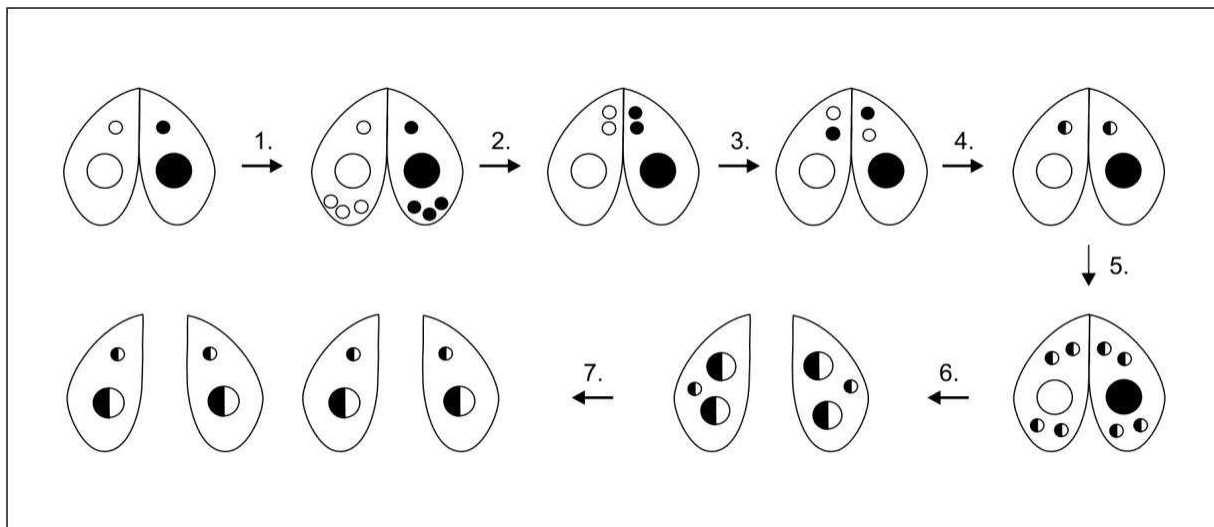


Abbildung 11: Konjugation bei *Tetrahymena* (Bockau 2006)

***Euglena gracilis* (Excavata)**

Euglena gracilis ist ein Mitglied der Euglenozoa. Diese stellen eine sehr distinkte Gruppe von Flagellaten dar, die in ihren drei Klassen Kinetoplastida, Diplonemida und Euglenida sowohl parasitäre, als auch heterotrophe (phagotrophe und osmotrophe) und photoautotrophe Vertreter aufweisen. Die zu den Euglenida gehörende *E. gracilis* (Abb. 12) ist eine phototrophe Art, wie die über sekundäre Endocytobiose erworbenen Chloroplasten mit Chlorophyll a und b zeigen.

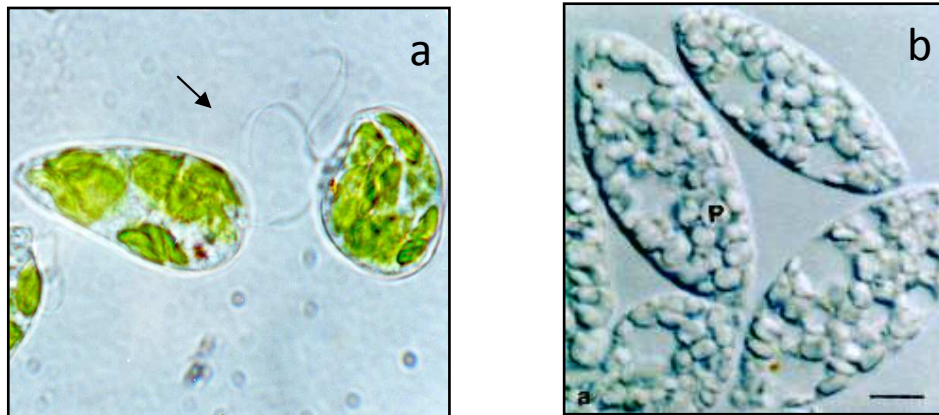


Abbildung 12: Lichtmikroskopische Aufnahme von *Euglena gracilis* a) im Licht kultiviert. b) im Dunkeln unter Zugabe von Kohlenstoffquellen kultiviert. P: Paramylon. 1 cm = 8 µm (Abbildung mit freundlicher Genehmigung von Prof'in Dr. A. Preisfeld, Bergische Universität Wuppertal)

Lediglich das exogene Vitamin B₁₂ muss in einer Kultur dem Medium zugeführt werden. Unter natürlichen Bedingungen sorgen Mikroorganismen für ausreichend viel Vitamin B₁₂ im umgebenden Medium (Süßwasser). Der Zellkörper von *E. gracilis* ist meist länglich zylindrisch, sie kann aber ihre Form über eine als Metabolie bezeichnete Fähigkeit verändern. Dabei verändert sie ihre Form durch Bewegung der unterhalb der Pellikula (Zellbegrenzung) liegenden Mikrotubuli und kann auch rundere Formen annehmen (Abb. 12 a). Normal bewegt sich *Euglena* mit der aus dem Reservoir herausragenden Geißel (die andere verbleibt darin) im Wasser fort (Pfeil Abb. 12 a). Sie kann sich aber auch „kriechend“ durch die metabolische Bewegung (*euglenid movement*) fortbewegen (Leedale 1967; 1978; Margulis & Chapman 2009). Die Pellikula ist das charakteristische morphologische Merkmal für die Euglenida (Preisfeld et al. 2001). Sie dient als äußere Zellbegrenzung und erscheint im Mikroskop streifig und helixartig verdreht. Im elektronenmikroskopischen Querschnitt kann man die longitudinale Anordnung der Streifen gut erkennen, die von Mikrotubuli unterlegt sind (Abb. 13). Schlechte Lebensbedingungen kann *Euglena* im unbeweglichen Palmella-Stadium wie eine Cyste in einer Gallerthülle ohne Geißeln überdauern. Bei günstigen Bedingungen werden die Geißeln in kurzer Zeit wieder hergestellt (Gruenberger 2007).

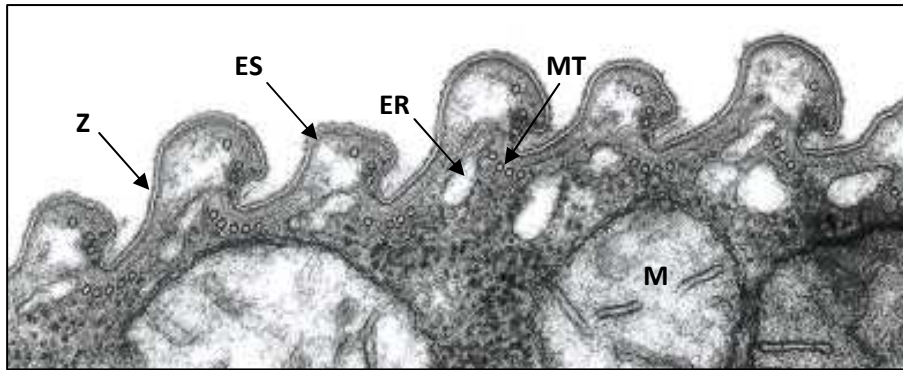


Abbildung 13: Transmissionselektronenmikroskopische Aufnahme der Pellikula von Euglenida.
Z: Zellmembran, **ES:** epiplasmatische Proteinschicht, **MT:** Mikrotubuli, **ER:** Endoplasmatisches Retikulum.
M: Mitochondrion. 1 cm = 0,2 µm (Abbildung mit freundlicher Genehmigung von Prof'in Dr. A. Preisfeld, Bergische Universität Wuppertal)

Die hier vorgestellten Protisten werden im Laborkurs „Kleine Einzeller ganz groß“ eingesetzt. Dieser wurde von Kaminska (2010) konzipiert und in vorliegender Studie mit minimalen Veränderungen durchgeführt und evaluiert.

1.1.4 Biowissenschaftliche Fragestellungen

Welche Bedingungen müssen erfüllt sein, um einen molekularen Identitätsnachweis in einem Experimentierkurs im Schülerlabor BeLL Bio erfolgversprechend einzusetzen?

Um diese Frage zu klären, ist es zunächst nötig, einen geeigneten Marker zu wählen. Da VNTR-Loci häufig für den Genetischen Fingerabdruck eingesetzt werden, sollten verschiedene DNA-Abschnitte auf ihre Eignung untersucht werden.

Für die Experimentierkurse sollte der beste Locus mit den optimierten Methoden ausgewählt werden.

Es sollte stichprobenhaft überprüft werden, ob die Sequenzbanden der VNTR-Loci sich unabhängig von ethnischer Zugehörigkeit nachweisen lassen.

Welche Anpassungen sind notwendig, um die Methoden und Arbeitsweisen zum Genetischen Fingerabdruck zum einen zeit- und kostensparend, zum anderen mit unbedenklichen Chemikalien bei hoher Reproduzierbarkeit durchführen zu können?

Hierzu sollten verschiedene Verfahren zur DNA-Extraktion aus bukkalen Schleimhautzellen optimiert werden. Anschließend sollten die PCR-Bedingungen so vereinheitlicht und vereinfacht werden, dass reproduzierbare Produkte als Bandenmuster entstehen. Weiterhin sollte ein weitgehend unbedenklicher, aber zuverlässiger Farbstoff zur Anfärbung der DNA im Agarose-Gel gefunden werden.

Unter welchen Bedingungen kann die Bestimmung der Pyruvatkonzentration in (künstlichem) Blut für das BeLL Bio experimentell umgesetzt werden?

Um den Schülern den Eindruck zu vermitteln mit – wenn auch künstlichen – Blutproben zu arbeiten, wurden verschiedene Lebensmittelfarbstoffe gesucht, welche es erlauben, die Pyruvatproben rot anzufärben. Dabei musste darauf geachtet werden, dass diese Farbstoffe nicht die enzymatische Umsetzung von Pyruvat zu Laktat im Photometer beeinflussen.

1.2 Fachdidaktischer Hintergrund

Der in Deutschland zu beobachtende Wandel hin zu einer forschungs- und wissensintensiveren Gesellschaft bringt einen wachsenden Bedarf an hoch qualifizierten Fachkräften der MINT-Fächer Mathematik, Informatik, Naturwissenschaften und Technik mit sich. Jährlich müssen etwa 58.000 in Ruhestand gehende Ingenieure, Mathematiker und Naturwissenschaftler durch MINT-Nachwuchs ersetzt werden (Hetze 2011). Diesen hohen gesellschaftlichen Bedarf können die derzeitigen Absolventenquoten in den naturwissenschaftlich-technischen Studiengängen nicht hinreichend decken (Huck et al. 2010). Anhaltend hohe Studienabbruchsquoten in den Natur- und Ingenieurwissenschaften gründen sich häufig in falschen Vorstellungen der Studierenden vom Studienfach oder von den zukünftigen beruflichen Tätigkeiten (Heubelein et al. 2008, Heubelein et al. 2009). Bereits auf Ebene der allgemeinbildenden Schulen ist eine MINT-Lücke zu beobachten. Immer mehr Schüler wenden sich von den Naturwissenschaften ab, was sich in der Wahl von Kursen in der Sekundarstufe II widerspiegelt (Prenzel & Ringelband 2001). Insbesondere in den naturwissenschaftlichen Fächern Chemie und Physik kommt nur eine relativ geringe Anzahl von Leistungskursen zustande (MSW NRW 2011b). Zwar ist diese Situation im Fach Biologie nicht derart ausgeprägt, jedoch werden auch diesen Schülern von den internationalen Schülervergleichsstudien der letzten Jahre deutliche Mängel in der naturwissenschaftlichen Grundbildung bescheinigt. Die Tragweite dieser Beobachtung wird bei näherer Inaugenscheinnahme des Begriffes der naturwissenschaftlichen Grundbildung bewusst. Dieser umfasst nicht nur das bloße Vorhandensein reproduktiven Faktenwissens. Er beschreibt vielmehr die Fähigkeit von Schülern, naturwissenschaftliches Wissen anzuwenden, naturwissenschaftliche Fragen zu erkennen und aus Belegen Schlussfolgerungen zu ziehen, um Entscheidungen zu verstehen und zu treffen, welche die natürliche Welt und die durch menschliches Handeln an ihr vorgenommenen Veränderungen betreffen (Deutsches PISA-Konsortium 2000). Naturwissenschaftliche Grundbildung ist somit eine wesentliche Voraussetzung für eine mündige Teilhabe am gesellschaftlichen Leben, eine reflektierte

Meinungsbildung und für Erfolg im naturwissenschaftlichen Berufsfeld. Sie trägt somit dazu bei, den Weg in die Wissensgesellschaft zu ebnen (Prenzel 2007, Dähnhardt, Haupt & Pawek 2009).

Schon für Grundschüler attestiert die *Trends in International Mathematics and Science Study* (TIMSS 2007), dass die Ziele des mathematisch-naturwissenschaftlichen Unterrichts in Deutschland nicht zufriedenstellend erreicht werden. Insbesondere die Problemlösekompetenz sowie die Fähigkeit des naturwissenschaftlichen Arbeitens werden in dieser Altersgruppe häufig nicht erlangt (Prenzel & Ringelband 2001). Viele Viertklässler verfügen zudem nur über elementare naturwissenschaftliche Kenntnisse und haben Schwierigkeiten, ihr Wissen produktiv anzuwenden (Bos et al. 2008). Euler (2001) führt Transferprobleme insbesondere darauf zurück, dass die Balance im derzeitigen Unterricht zu Gunsten der Fakten- und Wissensorientierung und zu Ungunsten der Praxisorientierung verschoben ist. Ein solcher Unterricht führt allenfalls zu „Know-what“- , aber nicht zu „Know-why“- oder „Know-how“-Ergebnissen (Euler 2001). Was die TIMSS-Grundschulstudie für den Primarbereich offenlegte, bestätigt das *Programme for International Student Assessment* (PISA, Baumert 2000) auch für Schüler gegen Ende ihrer Pflichtschulzeit. Mit dem im internationalen Vergleich unterdurchschnittlichen Abschneiden führte die erste PISA-Erhebung im Jahr 2000 eindrucksvoll das tatsächliche Ausmaß der Mängel vor Augen (Stanat et al. 2003). Der sogenannte „PISA-Schock“ setzte eine Diskussion um die Wertigkeit des deutschen Bildungssystems in Gang, die in vielfältigen Reformansätzen auf schulischer (Gundlach 2003) und außerschulischer Ebene mündete. Zur Qualitätsentwicklung schulischen Unterrichts wurden Programme wie SINUS, Biologie, Chemie oder Physik im Kontext eingerichtet. Auf außerschulischer Ebene entstanden eine Vielzahl von Einrichtungen, zu denen sich auch die Kinderuniversitäten, Schülerinfotage, Sommerunis und nicht zuletzt die Schülerlabore rechnen lassen (Dähnhardt, Haupt & Pawek 2009). Die Bemühungen um eine Verbesserung der Unterrichtsqualität schienen schnell Früchte zu tragen. Bewegte sich die naturwissenschaftliche Kompetenz der Schüler in Deutschland im Jahr 2000 noch unter dem Durchschnitt der PISA-Teilnehmerstaaten, so wurde nach der dritten PISA-Erhebung im Jahr 2006 bereits der „PISA-Jubel“ ausgerufen. Erstmals lagen hier die deutschen Schüler über dem internationalen Durchschnitt. Nicht zuletzt wurde dieser Erfolg dem gesteigerten Anwendungsbezug naturwissenschaftlichen Unterrichts, dem höheren Anteil eigenständigen Experimentierens sowie dem Einsatz didaktisch verbesserter Schülerexperimente zugeschrieben (Brandt et al. 2007, Prenzel 2007). Optimierungsbedarf besteht aber auch weiterhin. Selbst bei PISA 2009 ist die naturwissenschaftliche Kompetenz deutscher Schüler moderat ausgeprägt und liegt deutlich hinter dem Spitzenfeld Finnland, Japan und Korea. Mit 14,8 % erreicht ein immer noch zu großer Schüleranteil nur unzureichende Kompetenzen in den Naturwissenschaften. Diese Schüler sind im Alltag praktisch nicht in der Lage, einfache naturwissenschaftliche oder technische Probleme zu

verstehen und zu beurteilen. Ihre Chancen in eine Berufslaufbahn zu gelangen, die naturwissenschaftliches Verständnis erfordert, sind daher eher gering (Rönnebeck et al. 2010).

Das sich erst in jüngster Vergangenheit ausweitende Feld der Schülerlaborforschung sowie die Förderung der Schülerlaborszene von bildungspolitischer Seite (Ministerium für Innovation, Wissenschaft und Forschung NRW 2011) zeigen die großen Hoffnungen, die in diesen außerschulischen Lernort gesetzt werden. Inwieweit Schülerlabore durch die Möglichkeit des praktischen Experimentierens in authentischer Umgebung tatsächlich zu einer Verbesserung der naturwissenschaftlichen Grundbildung beitragen können, ist bisher allerdings noch unzureichend geklärt. Umfassende Evaluierungen stehen in den meisten Schülerlaboren immer noch aus (Engeln 2004).

Der folgende Abschnitt umreißt zunächst den Stand der vielfältigen Schülerlaborszene in Deutschland, um die Position des in dieser Studie untersuchten Lernortes BeLL Bio zu kennzeichnen. Im Anschluss daran wird das lerntheoretische Konzept des BeLL Bio vorgestellt und – unter Einbeziehung aktueller Schülerlaborforschung – werden lernrelevante Faktoren identifiziert, die den außerschulischen Lernort Schülerlabor kennzeichnen. Als zentrale naturwissenschaftliche Arbeitsweise kommt dem Experimentieren in Schülerlaboren eine besondere Rolle zu. Diese wird im Anschluss erläutert.

1.2.1 Lernort Schülerlabor

Die nicht einheitliche Verwendung des Begriffs Schülerlabor in der Literatur macht es zunächst notwendig, diesen für die Verwendung im folgenden Text näher einzugrenzen.

Als Schülerlabore gelten solche außerschulische Einrichtungen, in denen sich Schüler durch eigenes Experimentieren im Rahmen schulischer Veranstaltungen mit modernen Naturwissenschaften, Ingenieurwissenschaften, Mathematik und Informationstechnologie auseinandersetzen können. Dazu zählen Angebote in Universitäten, Fachhochschulen, Forschungseinrichtungen, Museen, Science Centern, Technologie- und Gründerzentren und der Industrie, die in dafür geeigneten Laboren Begegnungen mit diesen Themen ermöglichen. Die Angebote sind entweder regelmäßig, täglich oder wöchentlich zu besuchen (Dähnhardt et al. 2009).

In Bezug auf die beschriebenen Lernorte scheint diese Definition sehr weit gefasst. Neben den Experimentierangeboten an Forschungseinrichtungen und in der Industrie werden auch Museen und Science Center angesprochen. Diese Lernorte unterscheiden sich deutlich in ihrem formellen Charakter. Während das Museum in seiner klassischen Konzeption ein Ort der Selbstbildung und daher frei von äußeren Handlungszwängen ist (Parmentier 1995), spielt in anderen außerschulischen

Lernorten die Vermittlung von Wissen eine wichtige Rolle (Scharfenberg 2005, Berck & Graf 2010). Gropengießer und Kattmann (2008) ordnen Science Center hinsichtlich ihrer didaktischen Konzeption zwischen Museen und Schülerlaboren ein. Demnach ähnelt das Science Center in den Ausstellungsanteilen dem Museum. Viel stärker stehen im Science Center aber die praktische Erfahrung mit wissenschaftlichen Phänomenen und das positive Erlebnis im Vordergrund (Gropengießer und Kattmann 2008). Eine ähnliche Unterscheidung nimmt Brandt (2005) vor. Er differenziert zwischen Museen und Mitmachlaboren. Er sieht Science Center ähnlich den Museen, die unter motivationspsychologischen Kriterien gestaltet sind und über eine die Besucher überraschende, faszinierende, interessegeleitete und emotionale Ansprache zur Lernmotivation für die jeweiligen Exponate führen wollen (Brandt 2005). Die Mitmachlabore sieht er konzeptionell zwischen Science Center und Schule. Diese lösen die Schüler aus ihren gewohnten schulischen Strukturen heraus und stellen den Kontakt zu außerschulischen Institutionen und den dort arbeitenden Experten her (Brandt 2005). Im Grad ihres formellen Charakters lassen sich Schülerlabore somit auf einem Kontinuum zwischen dem eher informellen Science Center und dem Lernort Schule anordnen (Abb. 14). Durch die zum Teil sehr unterschiedlichen didaktischen Konzepte verschiedener Schülerlabore sind Variationen in ihrem formellen Charakter selbstverständlich möglich.

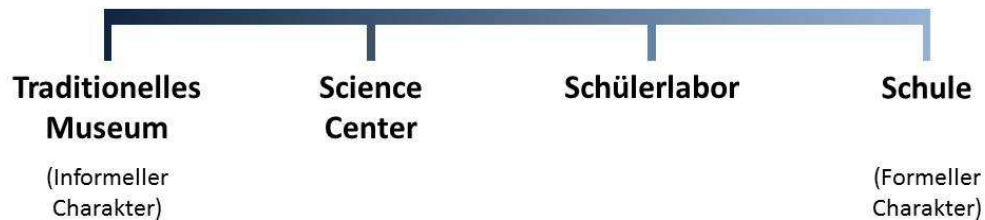


Abbildung 14: Das Schülerlabor BeLL Bio liegt in seiner konzeptionellen Offenheit auf einem Kontinuum zwischen dem formellen Charakter der Schule und dem eher informellen Charakter des traditionellen Museums (Abbildung in Anlehnung an Guderian 2007).

Während die anfänglich genannte Schülerlabor-Definition für den Lernort sehr weit greift, erscheint sie in Bezug auf die angesprochene Zielgruppe erweiterbar. Denn neben Veranstaltungen für Schüler, die sicherlich auch die Hauptzielgruppe dieser Labore ausmachen, werden in zunehmenden Maße Lehrerfortbildungen angeboten. Auch die Einbindung von Lehramtsstudierenden im Rahmen praxisorientierter Studien spielt eine bedeutende Rolle. Hinzu kommt, dass bereits vielerorts Schülerlaborkurse für die Öffentlichkeit zugänglich gemacht werden. Etwas offener gehalten ist in dieser Hinsicht die Definition der Initiative Zukunft durch Innovation (zdi):

Demnach sind zdi-Schülerlabore außerschulische Lernorte, an denen Schüler die Möglichkeit erhalten, unter nahezu professionellen Bedingungen technische Experimente durchzuführen oder naturwissenschaftlich zu forschen. zdi-Schülerlabore unterstützen weiterhin Schulen, Hochschulen,

andere Bildungsträger und Unternehmen bei der Förderung des MINT-Nachwuchses (Ministerium für Innovation, Wissenschaft und Forschung 2012).

Als derartige Unterstützung sind sicherlich auch die an den Schülerlaboren angebotenen Fortbildungen für Lehrkräfte zu sehen. Hier erhalten angehende und erfahrene Lehrer einen authentischen Einblick in die modernen Naturwissenschaften und erweitern ihr Methodenspektrum. Einen Überblick über die Bedeutung der Schülerlabore in Deutschland gibt der folgende Abschnitt.

1.2.1.1 Schülerlaborszene in Deutschland

Schüler erhalten in Deutschland bereits seit den siebziger Jahren die Möglichkeit, an außerschulischen Forschungseinrichtungen zu experimentieren und einen Einblick in die Arbeitswelt zu erhalten (<http://www.fortbildung.kit.edu/72.php>, abgerufen am 01.02.2012). Handelte es sich zunächst noch um vereinzelte Angebote, so entwickelte sich als Konsequenz der abnehmenden Attraktivität naturwissenschaftlicher Studienfächer und Berufe seit den neunziger Jahren eine deutlich vielfältigere außerschulische Bildungslandschaft (vgl. Euler 2009a/b). Nicht zuletzt angetrieben durch die Resultate der internationalen Schülervergleichsstudien nahm auch die Anzahl der außerschulischen Schülerlabore rasch zu, was häufig der persönlichen Initiative einzelner Wissenschaftler zu verdanken ist. Waren 1990 noch fünf Schülerlabore beim Bundesverband der Schülerlabore e.V. registriert, so sind es aktuell bereits dreihundert. Ein Großteil dieser Labore ist Universitäten oder außeruniversitären Forschungseinrichtungen angegliedert und bietet Laborkurse mit chemischen, physikalischen oder biologischen Inhalten an. Der Schwerpunkt des Kursangebotes richtet sich an Mittel- und Oberstufenschüler. Darüber hinaus existieren auch spezielle Angebote für Mädchen, Hochbegabte, Studierende des Lehramts, Referendare und Lehrer (Bundesverband der Schülerlabore e.V. 2012).

1.2.1.2 Schülerlabor-Netzwerke

Der Betrieb von Schülerlaboren ist mit hohen Personal- und Sachkosten sowie einem hohen zeitlichen Aufwand verbunden. Neben der Entwicklung und Durchführung von Kursen bestehen zentrale Aufgaben der Schülerlabore in der Akquirierung von Kursteilnehmern durch gezielte Werbemaßnahmen sowie in der Einwerbung von Personal- und Sachmitteln. Der Kontakt zu möglichen Förderern aus Wirtschaft und Politik verläuft insbesondere für einzelne Labore häufig wenig erfolgreich. Gerade diese ökonomischen Überlegungen sind es, die in den letzten Jahren ein Zusammenfinden einzelner Schülerlabore zu größeren Netzwerken bewirkt haben. Diese Kooperationen sind dringend notwendig, da bei geringerem Aufwand größere Wirkung erzielt

werden kann (Prenzel & Ringelband 2001). Die beschriebene Netzwerkbildung ist sowohl auf regionaler Ebene zu beobachten, wie z.B. im SaarLab (Hempelmann 2009), im ForschNet (Rückmann 2009), im SchulLaborBayern (Kiechle 2009), im GenaU (Detlefsen 2009), in den Bielefelder *teutolabs* und den Bergischen Science Labs in Wuppertal, als auch überregional, beispielsweise im Netzwerk der Schülerlabore in der Helmholtz-Gemeinschaft (Gatti 2009), im Netzwerk Genlabor & Schule (Nellen 2009), im Bundesverband Lernort-Labor e.V. (Bundesverband der Schülerlabore e.V., 2011) oder aber im Netzwerk der zdi-Zentren, zdi-Schülerlabore und zdi-RobertaZentren (Ministerium für Innovation, Wissenschaft und Forschung NRW 2012). In derartigen Netzwerken profitieren die einzelnen Labore von Synergieeffekten, die die Außenwahrnehmung erhöhen und das Auftreten gegenüber Politik und potentiellen Förderern erleichtern. Ein regelmäßiger Erfahrungsaustausch innerhalb der Netzwerke trägt zudem zur Qualität des Angebotes (Gatti 2009, Huck et al. 2010) bei.

1.2.1.3 Beschreibung des untersuchten Schülerlabors BeLL Bio

Das BeLL Bio ist als Teillabor der Bergischen Science Labs sowie als Mitglied des Schülerlabornetzwerkes der Initiative Zukunft durch Innovation (zdi) ausgezeichnet in die deutsche Schülerlaborszene integriert. Angegliedert ist das BeLL Bio an die Zoologie und Biologiedidaktik der Bergischen Universität Wuppertal. Mit einem Seminarraum, zwei Vorbereitungsräumen und einem großzügigen, modernen Laborraum können im BeLL Bio auch große Schulklassen mit bis zu dreißig Schülern experimentieren. Das Kursangebot richtet sich vorwiegend an Biologiekurse der Sekundarstufe I und II. Aktuell werden deutschsprachig die Laborkurse „Genetischer Fingerabdruck – Dem Täter auf der Spur“, „Außer Atem – Der Laktatdehydrogenase auf der Spur“, „Kleine Einzeller ganz groß“, „CrimeLab“ und „Reine Nervensache“ angeboten, von denen die ersten drei Untersuchungsgegenstand der vorliegenden Studie sind. Darüber hinaus erhalten Schüler in speziellen bilingualen Kursen die Möglichkeit, in die Nutzung des Englischen als Arbeitssprache hineinzuwachsen. Das curriculumsnahe Kursangebot des BeLL Bio knüpft unmittelbar an den schulischen Unterricht an, so dass die unterrichtliche Vorbereitung integrales Element jedes Laborkurses ist. Im BeLL Bio sollen Schüler die Möglichkeit erhalten, Versuche in der authentischen Lernumgebung Labor durchzuführen, die so im schulischen Kontext nicht umsetzbar wären. Erste Fortbildungen für Referendare zur Thematik „Einbindung von BeLL Bio-Besuchen in den schulischen Unterricht“ wurden bereits durchgeführt. Fortbildungen für erfahrene Lehrkräfte sind in Planung.

In der Formulierung seiner Ziele orientiert sich das BeLL Bio an den von Engeln und Euler (2004) sowie Dähnhardt et al. (2009) identifizierten Absichten von Schülerlaboren, erweitert und spezifiziert diese aber um einige Aspekte. Das BeLL Bio möchte

- Interesse für die Naturwissenschaft Biologie wecken.
- ein spannendes, nachhaltiges Lernerlebnis bieten.
- Selbstvertrauen im Umgang mit biologischen Sachverhalten bewirken.
- zur Auseinandersetzung mit biologischen Sachverhalten aktivieren.
- authentische Einblicke in biologisch-technische Berufe und Forschung ermöglichen und Interesse für das entsprechende Berufsfeld wecken.
- ein zeitgemäßes Bild der modernen Biologie und ihrer Bedeutung für die Gesellschaft vermitteln (vgl. Engeln & Euler 2004, vgl. Dähnhardt et al. 2009).

Wie an den oben genannten Formulierungen deutlich wird, steht im BeLL Bio nicht der reine Eventcharakter im Vordergrund. Neben einer affektiven Aktivierung möchte das BeLL Bio in erster Linie auch einen nachhaltigen kognitiven Lernerfolg ermöglichen. Insbesondere Schülern, die ein naturwissenschaftliches Studium anstreben, möchte das BeLL Bio darüber hinaus im Sinne wissenschaftspropädeutischen Arbeitens ein Methodenbewusstsein vermitteln, wie es auch in den Richtlinien und Lehrplänen für die gymnasiale Oberstufe (MSWWF NRW 1999) gefordert wird. Als indirektes Ziel im Sinne Behrendts (2009) wird auch die Ideengabe für erfahrene, aber auch für angehende Lehrkräfte gesehen. Insbesondere angehende Biologielehrer erhalten bereits im Rahmen ihres Studiums die Möglichkeit, an der Entwicklung, Vorbereitung, Durchführung und Evaluierung von BeLL Bio-Kursen mitzuwirken. Neben einer Weiterentwicklung ihrer eigenen Methodenkenntnis bietet das BeLL Bio zusätzliche außerschulische Erfahrung im Umgang mit Schülern. Darüber hinaus erhalten Studierende Einblick in biologiedidaktische Forschungsmethodik.

Ob diese hochgesteckten Ziele, wie sie in ähnlicher Form auch in anderen Schülerlaboren formuliert werden, im BeLL Bio tatsächlich erreicht werden, soll die vorliegende Studie zeigen. Da der Nutzen von Laborkursen häufig – gerade in Anbetracht der relativ hohen Kosten – sehr kritisch gesehen wird (Borrmann 2008), erscheinen derartige Untersuchungen umso wichtiger. Im folgenden Abschnitt wird – basierend auf dem lernpsychologischen Ansatz des gemäßigten Konstruktivismus – das lerntheoretische Konzept der Lernumgebung BeLL Bio vorgestellt.

1.2.1.4 Lerntheoretisches Konzept des BeLL Bio

Die lerntheoretische Basis des BeLL Bio-Konzeptes bildet der gemäßigte Konstruktivismus nach Reinmann und Mandl (2006). Als Lehr-Lerntheorie mit hoher Praxisrelevanz vereint er die Grundideen des Instructional Designs (ID) mit denen des Konstruktivismus. Beide Positionen sollen daher kurz skizziert werden. Die kleinschrittige systematische Planung und die Durchführung strikter Lernerfolgskontrollen sind typische Merkmale gegenstandszentrierten Unterrichts im Sinne des ID.

Wissen wird hier als Gegenstand betrachtet, der unter Einhaltung bestimmter Regeln vom Lehrer auf den Lerner übertragen werden kann. ID-Ansätze bemühen sich, Instruktionspläne zu erstellen, die dem Lehrenden vorgeben, unter welchen Voraussetzungen er welche Lehrmethoden einsetzen soll. Systematische Lernerfolgskontrollen werden im ID eingesetzt, um den Lernfortschritt und damit die Differenz zum gewünschten Endzustand zu ermitteln. Damit steht das ID in der Tradition des Behaviourismus, welcher menschliches Lernen auf wenige zentrale Lernmechanismen zurückführt (Reinmann & Mandl 2006).

Im Gegensatz zur der zuvor beschriebenen technologischen Position betrachten konstruktivistische Vertreter Lernen nicht als passiven, sondern als einen auf konstruktiver Eigenaktivität basierenden Prozess. Der Erkenntnistheorie des radikalen Konstruktivismus zufolge existiert eine objektive Realität nicht. Vielmehr basiert demnach die gesamte menschliche Wahrnehmung auf Konstruktion und Interpretation, so dass die Wirklichkeit ein subjektiv konstruiertes Phänomen bleibt (Glaserfeld 2010, Watzlawick 2011). Vor diesem Hintergrund würden sich die Gestaltung einer Lernumgebung sowie Vermittlungsversuche einer Lehrkraft erübrigen, da jeder Schüler den Lerngegenstand anders rezipieren würde. Der neue Konstruktivismus nimmt eine deutlich moderatere Position ein. Gemeinsames Wissen ist demnach durchaus realisierbar, da es im Rahmen sozialer Interaktionen entwickelt und ausgetauscht wird. Lernen wird hier als aktiv-konstruktiver Prozess betrachtet, der stets in einem bestimmten Kontext erfolgt. Um den Anwendungsbezug des Wissens zu sichern und träges Wissen zu vermeiden, sollten die Lernprozesse möglichst authentische, bedeutungshaltige Probleme beinhalten. Dem Aspekt der Problemorientierung wird im gemäßigt konstruktivistischen Ansatz besondere Bedeutung beigemessen. Ausgangspunkt des Lernens sollten demnach authentische Problemstellungen sein, die aufgrund ihres Alltagsbezugs und ihrer Relevanz zur aktiven Auseinandersetzung mit dem Problem motivieren. Die Arbeit an der Problemlösung sollte kooperativ in Gruppen und mit instruktionaler Unterstützung erfolgen. Diese Balance aus Instruktion und Konstruktion soll die Überforderung von Schülern mit eher ungünstigen Lernvoraussetzungen vermeiden und leistungstärkeren Lernern die Möglichkeit der Selbststeuerung des Lernprozesses eröffnen (Reinmann & Mandl 2006).

Engeln und Euler (2004) gehen davon aus, dass Schülerlabore viele derjenigen Kriterien erfüllen, die unter lerntheoretischer Perspektive als relevant für eine aktive Wissenskonstruktion durch die Lernenden angesehen werden. Da konstruktivistische Lernumgebungen besondere räumliche Bedingungen erfordern, können viele Schulen, die noch eher am traditionell-instruktiven Lernverständnis ausgerichtet sind, diese Voraussetzungen nicht erfüllen (Brandt 2005). Gegen den konstruktivistischen Charakter von Schülerlaborexperimenten spricht das oft angeführte Argument des Kochbuchcharakters derartiger Versuche (Hofstein & Lunetta 2003). Entsprechende Experimente können nach Anleitung abgearbeitet werden, ohne die Schritte im Sinne konstruktiver Eigenleistung

reflektieren zu müssen. Auch wenn in vielen Schülerlaboren – so auch im BeLL Bio – skriptgeleitete Versuche dominieren, so fordern die im BeLL Bio durchgeführten, relativ komplexen Experimente dennoch ständige Selbstorganisation. In Kleingruppen müssen die Schüler beispielsweise über die Wahl der korrekten Arbeitsgeräte entscheiden, mit Mitschülern oder mit dem betreuenden Kurspersonal über deren Bedienung und Verständnisprobleme diskutieren, sich die Zeiträume – beispielsweise während des Zentrifugierens oder während photometrischer Messungen – ökonomisch einteilen und ihre Beobachtungen reflektieren. Tabelle 1 gibt einen Überblick über fünf Leitlinien, welche es nach Reinmann und Mandl (2006) bei der Gestaltung einer nach gemäßigten konstruktivistischen Prinzipien gestalteten Lernumgebung zu beachten gilt. Ferner wird deren Realisierung in den Laborkursen des BeLL Bio aufgezeigt.

Tabelle 1: Leitlinien für problemorientierten Unterricht und deren Bedeutung nach Reinmann & Mandl (2006) und Realisierung in den Laborkursen des BeLL Bio

Leitlinie	Bedeutung	Realisierung im BeLL Bio
Situiert und anhand authentischer Probleme lernen	Die Lernumgebung ist so zu gestalten, dass sie den Umgang mit realistischen Problemen und authentischen Situationen ermöglicht und anregt.	Die Schüler lernen in einem authentischen Laborraum, welcher der Universität angegliedert ist unter Einsatz von Laborgeräten, wie sie in der Forschung und in labortechnischen Berufen eingesetzt werden. Sie werden von Wissenschaftlern betreut und erhalten in den Laborkursen Problemstellungen, welche es zu lösen gilt (z.B. Welcher Tatverdächtige kann der Mörder sein?).
In multiplen Kontexten lernen	Die Lernumgebung ist so zu gestalten, dass das Gelernte auch auf andere Problemstellungen oder Fallbeispiele übertragen werden kann.	Im Rahmen der Nachbesprechung jedes BeLL Bio-Kurses werden weitere Anwendungsmöglichkeiten der im Labor erlernten Methoden reflektiert und Einblicke in das Berufsfeld gegeben, welches von eben diesen Methoden Gebrauch macht.
Unter multiplen Perspektiven lernen	Beim Lernen sollte berücksichtigt werden, dass man einzelne Inhalte oder Probleme aus verschiedenen Blickwinkeln sehen oder unter verschiedenen Aspekten beleuchten kann.	Wie in jeder Lernumgebung sehen die Schüler auch im Schülerlabor Inhalte und Probleme zunächst aus ihrer eigenen Perspektive. Die kursbegleitenden Betreuer geben Erklärungen aus ihrer Perspektive bzw. aus der Perspektive der Wissenschaft.
In einem sozialen Kontext lernen	Die Lernumgebung ist so zu gestalten, dass sie kooperatives Lernen und Problemlösen in Gruppen ermöglicht und fördert.	Die Schüler arbeiten in der Regel in Zweierteams. Dies sichert hohe Eigenaktivität bei ständigem gegenseitigem Austausch. Ökonomisches Arbeiten erfordert zudem in vielen Arbeitsschritten die Absprache mit anderen Zweierteams. Weiterhin stehen die Schüler in ständigem Austausch mit den Kursbetreuern.
Mit instruktionaler Unterstützung lernen	Die Lernumgebung ist so zu gestalten, dass neben vielfältigen Möglichkeiten eigenständigen Lernens in komplexen Situationen auch das zur Bearbeitung von Problemen erforderliche Wissen bereitgestellt wird.	In der komplexen Laborumgebung werden die Schüler geleitet durch das Versuchsskript sowie durch die kursbegleitenden Betreuer. Diese nehmen eine eher beratende Position ein.

Die Kombination aus hoher Eigenaktivität und skriptgeleiteter Instruktion im BeLL Bio kommt damit sicherlich der Intention des gemäßigten Konstruktivismus entgegen, wenn auch in einer Form moderater Realisierung. Eine solche Umsetzung scheint allerdings erfolgversprechender zu sein als eine gänzlich offene Lernumgebung (Euler 2009b, Parchmann 2009).

Wie vorangegangene Schülerlaborstudien zeigen (z.B. Engeln 2004, Brandt 2005, Pawek 2009), hängt die Wirkung der Lernumgebung Schülerlabor sowohl von den Merkmalen der Lernumgebung selbst als auch von der persönlichen Disposition der Lernenden ab. Die Merkmale, welche die Persönlichkeit der Schüler ausmachen, werden im Folgenden als Personenvariablen, die Merkmale der Lernumgebung als Laborvariablen bezeichnet. Die Bedeutungen dieser Variablen im Kontext Schülerlabor werden in den folgenden Abschnitten unter Einbezug aktueller Schülerlaborforschung vorgestellt.

1.2.2 Laborvariablen

In Orientierung an Engeln (2004), Brandt (2005) und Pawek (2009) erscheinen als beachtenswerte Einflussgrößen der Lernumgebung Schülerlabor der Alltagsbezug, die Authentizität, die Person des Kursbetreuers, die Transparenz der Versuchsanleitung, die Zusammenarbeit mit Mitschülern, die Offenheit der Lernumgebung, der empfundene Leistungsdruck sowie die Laboratmosphäre. Die Bedeutung dieser Laborvariablen wird nachfolgend differenzierter betrachtet.

1.2.2.1 Alltagsbezug

Wie Barke und Harsch (2011) feststellen, haben Jugendliche ein natürliches Interesse daran, mehr über die Gegenstände und Vorgänge ihrer Lebenswelt zu erfahren. An diese Neugier kann im Unterricht angeknüpft und die Fragestellungen aus dem Alltag altersgemäß aufbereitet werden. Entsprechende Forderungen finden sich auch in aktuellen Lehrplänen. Diese zielen darauf ab, durch die Anknüpfung an Alltagserfahrungen und Vorstellungen der Schüler an naturwissenschaftliche Konzepte, Sicht- und Arbeitsweisen heranzuführen. Vorkenntnisse können so aufgegriffen und in Richtung tragfähiger fachlicher Konzepte gesteuert werden (MSW 2008b). Zudem bildet die Berücksichtigung der subjektiven Alltagsvorstellungen einen erfolgversprechenden Ausgangspunkt für selbstbestimmtes Lernen (Gebhard 2008). So stellen Rakoczy et al. (2008) fest, dass der Alltagsbezug, den Schüler der achten und neunten Klasse im Mathematikunterricht wahrnehmen, deutlich zur Entwicklung selbstbestimmter Motivation beiträgt. Alltagsbezüge leisten einen deutlichen Beitrag zum Verständnis, was sich auch für Schüler mit geringeren fachlichen Vorkenntnissen im Kontext des Schülerlabors als bedeutsam erweisen könnte. Insbesondere im

Rahmen von Unterrichtsexperimenten wird die Chance gesehen, Alltagsvorstellungen der Schüler aufzugreifen und an Problemsituationen aus ihrer Lebenswelt anzuknüpfen (Hopf und Wiesner 2003, Scharfenberg 2005). Diesen Forderungen kommt man in vielen Schülerlaboren nach. Hier werden Versuche in der Regel in für Schüler interessante thematische und methodische Kontexte eingebettet (Euler 2009a/b). Diese schließen an das aus Alltagserfahrungen herrührende Vorwissen der Schüler an und kommen damit den Vorstellungen des gemäßigten Konstruktivismus nach einer situierten Lernumgebung und aktiver Wissenskonstruktion entgegen (vgl. Punkt 1.2.1.4.). Entsprechend geht Roth (2009) davon aus, dass auch der Lernerfolg vom spezifischen Lehr-/Lernkontext abhängt. Einen positiven Zusammenhang von Alltagsbezug und Lernerfolg können Nick und Urhahne (2004) im Rahmen einer Vergleichsstudie feststellen. Sie gehen davon aus, dass der Alltagsbezug zu einer gesteigerten Motivation der Lernenden führt und auf diese Weise die Lernbereitschaft gegenüber komplexeren, fachspezifischeren Inhalten gesteigert wird. Entsprechend bemühen sich auch die BeLL Bio-Kurse um die Gewährleistung von Alltagsbezügen.

1.2.2.2 Authentizität

Schülerlabore haben das Potential, den schulischen Unterricht durch authentische Erfahrungen zu ergänzen (Biermann et al. 2009, Euler 2009a/b). In Anlehnung an Scharfenberg (2005) wird in der vorliegenden Studie unter Authentizität eine möglichst große Annäherung an die reale Forschungswelt verstanden. Pawek (2009) identifizierte die Authentizität des Schülerlabors als einen zentralen Faktor für die Generierung aktuellen Interesses und die Steigerung des Fähigkeitsselbstkonzeptes. Auch Helms (1998) stellt fest, dass neben der Offenheit insbesondere die Authentizität der Fragestellungen eine naturwissenschaftliche Aufgabenstellung für Schüler interessant macht. Authentische Projekte geben Lernern die Möglichkeit, ihr Vorwissen in neuen Situationen anzuwenden. Erfahrungen im Klassenraum führen dagegen zu Langeweile und Demotivation (Helms 1998). Die im BeLL Bio durchgeführten Kurse bieten Schülern Authentizität durch Aufenthalt und praktisches Experimentieren in einem universitären, forschungsnahen Labor. Im kommunikativen Austausch mit den kursbetreuenden Wissenschaftlern sowie Studierenden, erhalten die Kursteilnehmer authentische Einblicke in reale Forschungs- und Laborarbeit, Unialltag und Studium. Letztendlich runden auch die praktische Auseinandersetzung mit modernen Laborgeräten sowie das Tragen der labortypischen Schutzausstattung (Kittel, Schutzbrille, Handschuhe) den authentischen Eindruck ab.

1.2.2.3 Kursbetreuer

Auch wenn die Kursbetreuer in der Rolle der Experten bereits als authentisches „Merkmal“ des Lernortes Schülerlabor Erwähnung fanden, so kommt ihnen im Zuge des Lehr-/Lernprozesses weitere Bedeutung zu. In Anlehnung an Hascher und Mosers (2001) Beschreibung eines Praktikumslehrers übernehmen die Kursbetreuer im BeLL Bio die Funktion, den Schülern als Berater bei Problemen zur Seite zu stehen und zur Reflexion des eigenen Handelns anzuregen. Im Unterschied zur Lehrperson im schulischen Unterricht, die gleichzeitig viele, meist sehr heterogene Schülergruppen betreut, steht im BeLL Bio ein Betreuer für etwa sechs Schüler zur Verfügung. Er kann daher im Sinne guter naturwissenschaftlicher Lehre (Cimer 2007) sehr viel direkter auf Präkonzepte der Kursteilnehmer eingehen, die Aufmerksamkeit auf bestimmte Aspekte fokussieren, zur Kommunikation anregen und positives Feedback geben. So eröffnen die Kursbetreuer den Schülern ausreichend Gelegenheit zum selbstständigen Agieren und die notwendige Anleitung, um Überforderungen zu vermeiden. Sie gewährleisten damit die Balance aus Instruktion und Konstruktion und ermöglichen soziale Interaktion, wie es für die Gestaltung gemäßigt konstruktivistischer Lernumgebungen gefordert wird (vgl. Punkt 1.2.1.4.). Nicht zuletzt ist es Aufgabe der Kursbetreuer, die Einhaltung der Sicherheit im Labor zu gewährleisten, denn wissenschaftliches Arbeiten verlangt die Berücksichtigung bestimmter Regeln (Masters & Nott 1998).

1.2.2.4 Transparenz der Versuchsanleitung

Bietet die Person des Kursbetreuers Beratung „auf Nachfrage“, so stellen die Versuchsanleitungen ein kontinuierliches instruktionales Element vieler Schülerlaborkurse dar. Entsprechend wird von Versuchsanleitungen gefordert, dass diese klar und eindeutig sind (Klautke 1974). Auch Aufdermauer und Hesse (2006a) nennen die Verständlichkeit sowie die Vollständigkeit als ein Qualitätskriterium von Experimentieranleitungen. Die Anleitungen sollten demnach eindeutig und einfach sowie motivierend formuliert sein. Unbekannt erscheinende Fachbegriffe und Materialien müssen erläutert werden, alle benötigten Materialien und Einzelschritte erwähnt werden. Veranschaulichend wirkt der Einsatz von Schemata oder Abbildungen (Aufdermauer & Hesse 2006a). Weiterhin ist es von Bedeutung, auch komplexe Versuchsanleitungen nicht auf reinen Kochbuch-Charakter zu reduzieren. Bei einem reinen Abarbeiten von Arbeitsanweisungen besteht die Gefahr, allenfalls manuelle Fertigkeiten, nicht aber systematisches Denken zu schulen (Engeln & Euler 2004). Nur die Balance aus gelungener Anleitung und eigener Denkleistung führt zu einer nachhaltigen Erkenntnisgewinnung (Parchmann 2009). Aus diesem Grund enthalten die im BeLL Bio eingesetzten Versuchsanleitungen Elemente, die die Eigenaktivität und Kooperation unterstützen. Dies können Aufgaben oder

Informationsfelder sein, die zur Reflexion der getätigten Arbeitsschritte beitragen oder Hinweise auf die Notwendigkeit zur Kooperation mit anderen Arbeitsgruppen geben.

1.2.2.5 Zusammenarbeit mit Mitschülern

Innovationen entstehen heute insbesondere durch Zusammenarbeit von Wissenschaftlern aus verschiedenen Disziplinen. „So genannte Social Skills wie Team- und Kommunikationsfähigkeit sind dazu neben der fachlichen Kompetenz unabdingbar.“ (Dähnhardt et al. 2009). Zudem wird das Empfinden von sozialer Eingebundenheit, neben dem Erleben von Autonomie und Kompetenz, als wichtige Voraussetzung für die Entstehung intrinsisch motivierten Verhaltens angesehen (Deci & Ryan 2000). Auch im Schülerlabor BeLL Bio ist deshalb Zusammenarbeit mit den Mitschülern gefragt. Die erfolgreiche Durchführung der Versuche verlangt Absprachen mit dem Gruppenpartner oder benachbarten Arbeitsgruppen. Zentrifugierdauern müssen abgestimmt, Photometer-, Mikroskop- oder PC-Belegungen koordiniert und Ergebnisse verglichen werden. Sachbezogene Fragestellungen á la „Sag mal, wie habt ihr die Pipette eingestellt?“ oder „Wie habt ihr das ausgerechnet?“ werden ausgetauscht und diskutiert. Wie Engeln (2006) feststellt, ist für Mädchen die Zusammenarbeit beim Experimentieren von größerer Bedeutung als für Jungen. Einen Einfluss der Variable Zusammenarbeit auf das aktuelle Interesse kann sie nicht beobachten (Engeln 2004).

1.2.2.6 Offenheit

Eine offene Lernumgebung zeichnet sich durch einen geringen Grad an Strukturierung und weitreichenden Möglichkeiten zum selbstgesteuerten Lernen aus. Autonomieempfinden wird neben Kompetenzerleben und der Erfahrung sozialer Eingebundenheit als zentrale Voraussetzung für die Entstehung intrinsischer Motivation gesehen (Deci & Ryan 2000, Schiefele & Streblow 2006). Intrinsische Motivation zeigt sich in persönlichem Interesse einer Person und führt zu interessenorientierten Handlungen (Gropengießer & Kattmann 2008). Der institutionelle Rahmen des schulischen Unterrichts bietet Schülern kaum Möglichkeiten zur Selbstbestimmung, was das Entstehen intrinsischer Motivation verhindert (Brandt 2005). Wie in der Abgrenzung des Schülerlabors von den Institutionen Schule, Science Center und Museum dargestellt (vgl. Punkt 1.2.1.), lässt sich der formelle Charakter der meisten Schülerlabore auf einem Kontinuum zwischen Science Center und Schule abbilden. Die individuelle Position eines Schülerlaborkurses hängt unter anderem auch davon ab, welche Freiheiten den Schülern bei der Auseinandersetzung mit den fachlichen Gegenständen gelassen werden. Sunal et al. (2008) differenzieren zwischen vier Laborkurstypen in Abhängigkeit vom Grad ihrer Offenheit (Tabelle 2).

Tabelle 2: Rolle der Schüler sowie der Lehrkraft in den vier Laborkurstypen (in Anlehnung an Sunal et al. 2008; SuS=Schülerinnen und Schüler)

	Laborkurstyp	Schülerrolle	Lehrerrolle
<div style="display: flex; align-items: center;"> <div style="writing-mode: vertical-rl; transform: rotate(180deg); font-weight: bold; margin-right: 10px;">Grad der Offenheit</div> <div style="background: linear-gradient(to bottom, #e0e0ff, #000080); width: 20px; height: 300px; margin-left: 10px;"></div> </div>	Bestätigendes Experimentieren <i>Confirmation</i>	Versuchsergebnisse sind bereits bekannt und sollen nur noch einmal reproduziert werden; SuS achten darauf, was von der Lehrperson vorgegeben wird und führen Tätigkeiten entsprechend der Vorgaben aus.	Liefert Fragestellungen, Anleitungen, Materialien sowie die erwarteten Ergebnisse und bestimmt über die Eignung der Schülerantworten
	Strukturiertes Forschen <i>Structured inquiry</i>	Führen vorgegebene Versuche durch, konstruieren selbst Antworten und schließen auf deren Geeignetheit zur Beantwortung der Fragestellung	Liefert Fragestellungen, Anleitungen und Materialien, um die SuS selbst arbeiten zu lassen und unterstützen SuS in ihrer Rolle
	Geleitetes Forschen <i>Guided inquiry</i>	Planen Versuche, führen diese durch, konstruieren Antworten und schließen auf deren Geeignetheit zur Beantwortung der Fragestellung	Liefert Fragestellungen und Materialien, um die SuS selbst arbeiten zu lassen und unterstützt die SuS in ihrer Rolle
	Offenes Forschen <i>Open inquiry</i>	Legen die Fragestellungen selbst fest, planen Versuche, führen diese durch, konstruieren Antworten und schließen auf deren Geeignetheit zur Beantwortung ihrer Fragestellung	Unterstützt SuS in ihrer Rolle
Geringe Offenheit			
Hohe Offenheit			

Sunal et al. (2008) betonen, dass jeder der vier Laborkurstypen seine Berechtigung hat. Während das bestätigende Experiment insbesondere beim Erlernen neuer Labormethoden geeignet erscheint, fördert größere Offenheit insbesondere höhere Qualitäten des Denkens. Entsprechend beobachtet Zehren (2009), dass der mehrfache Besuch von Experimentierkursen hoher Offenheit dazu führt, dass Schüler fachlich weiter in die Tiefe gehende Fragestellungen entwickeln. Hopf und Wiesner (2003) stellen in offener gehaltenen Experimentierkursen eine Zunahme von Interesse, motivationalem Erleben, Selbstwirksamkeitserwartung und Kompetenzerleben von Schülern fest. Engeln (2004) beobachtet dagegen kaum einen Einfluss der Offenheit auf das Interesse. Sie geht zudem davon aus, dass auch eine geringe Offenheit zu kognitiver Aktivierung führen kann. Unter Berücksichtigung der unterschiedlichen praktischen Vorerfahrung könnte ein zu hoher Grad an Offenheit auch mit dem Problem einhergehen, dass nur wenige Kursteilnehmer praktisch arbeiten und andere nur bei der Arbeit zusehen (Davies et al. 2008). Gerade weil das Erlernen neuer Methoden eine unabdingbare Voraussetzung für eigene forschende Tätigkeit darstellt, kann es ebenfalls die wahrgenommene Selbstständigkeit fördern (Messner 2009). Mit zunehmenden Fertigkeiten sollten den Schülern dann zunehmend Möglichkeiten zum selbstständigen Arbeiten gegeben werden (Howe & Smith 1998). Insbesondere komplexe Versuche mit hohem wissenschaftlichem Anspruch, wie sie häufig in der Oberstufe von Interesse sind, sind zudem kaum ohne detaillierte Anleitung in Verbindung mit intensiver Betreuung behandelbar (Zehren 2009). Aus

Sicherheitsgründen wäre es außerdem unverantwortlich, die Schüler nach Versuch und Irrtum experimentieren zu lassen (Masters & Nott 1998). Im BeLL Bio erhalten die Schüler daher zunächst eine Problemstellung mit Einweisung in die zur Problemlösung benötigten Laborgeräte einschließlich einer Belehrung über sicheres Arbeiten im Labor. Diese Vorbereitung befähigt sie, im Anschluss relativ selbstständig – geleitet durch eine Versuchsanleitung – den Problemlöseweg zu beschreiten, unterstützt von den Kursbetreuern, die sie auf diesem Weg begleiten. Damit lässt sich das Konzept des BeLL Bio im Grad seiner Offenheit im Laborkurstyp des strukturierten Forschens nach Sunal et al. (2008, vgl. Tab. 2) wiederfinden.

1.2.2.7 Druck

Die Offenheit einer Lernumgebung kann Einfluss nehmen auf den empfundenen Druck. Der Leistungsabfall unter Druck (*choking under pressure*) ist ein in der Psychologie beschriebenes Phänomen, welches beinhaltet, dass sich die Leistung einer Person bei einer zu starken Erhöhung des psychischen Drucks in der Regel verschlechtert. Als Erklärung für diese Beobachtung wird angenommen, dass Druck die Person vom eigentlichen Gegenstand ihrer Aktivität ablenkt, sei es, weil sie ihre Aufmerksamkeit auf andere Aktivitäten oder aber auf ihre eigene innere Wahrnehmung richtet (Lewis & Linder 1997). Entsprechend verwundert es nicht, dass Druck auch als negativer Prädiktor intrinsischer Motivation gesehen wird (Deci & Ryan 2012). Die *Cognitive Load Theory* (CLT) nach Chandler und Sweller (1991) beschäftigt sich spezieller mit diesem Phänomen auf kognitiver Ebene. Sie beschreibt die Gefahr, dass die kognitive Aufmerksamkeit des Lernalers bei anspruchsvollen Problemlösetätigkeiten vom eigentlichen Lerninhalt abgelenkt wird. Den Beitrag instruktorischer Maßnahmen erkennt die CLT darin, dass diese bei angemessener Aufbereitung zu einer kognitiven Entlastung des Lernalers führen und die kognitiven Ressourcen des Lernalers auf den Lerngegenstand fokussieren können (Chandler & Sweller 1991).

Schüler haben häufig relativ wenig Erfahrung im praktischen Experimentieren (Scharfenberg 2005). So ist anzunehmen, dass sie sich insbesondere bei sehr offen gehaltenen Experimentalaufgaben schnell überfordert fühlen könnten (Kirchner & Priemer 2007). Eine übersichtliche Gestaltung des Laborkurstages, transparente Versuchsskripte und unterstützende Betreuer können dazu beitragen, den Druck zu reduzieren und die Aufmerksamkeit der Schüler auf wesentliche Versuchsinhalte zu fokussieren. Ein souveränes Auftreten der Kursbetreuer sowie das Gefühl, ausreichend Zeit zu haben und keine Angst vor Fehlern haben zu müssen, lassen Schüler die Durchführung der Versuche stressfrei erleben (vgl. Hillebrandt & Dähnhardt 2005, vgl. Pfangert-Becker 2010).

1.2.2.8 Atmosphäre

„Ein Raum wirkt stärker als tausend Worte und Bilder, weil er den Menschen ganzheitlich einnimmt und einstimmt.“ (Pfister 2008). Diese Aussage trifft sicherlich genauso für den Biologieraum in der Schule, als auch für den Laborraum im Schülerlabor zu. Zwar gibt es wenige Schulen mit sehr gut ausgestatteten, laborähnlichen Biologieräumen und auf der anderen Seite Schülerlabore, in denen die Schüler unter seminarraumähnlichen Bedingungen experimentieren, in der Regel unterscheiden sich die räumlichen Bedingungen in Schülerlabor und Schule aber deutlich voneinander. In den meisten Fällen ist daher die Situation Laborraum für viele Schüler neu. Im Falle des BeLL Bio arbeiten die Schüler in einem sehr modernen, gut ausgestatteten, annähernd tageslichthellen Raum. Doch die im Laborraum empfundene Ortsatmosphäre ergibt sich nicht nur aus diesen gestalterischen Merkmalen des Laborraumes, sondern wird ebenso beeinflusst durch die Atmosphäre der Situation. Diese ergibt sich unter anderem aus den sozialen Interaktionen der vor Ort anwesenden Personen (Pfister 2011), aber auch durch die Angliederung des BeLL Bio an die Universität. Die universitäre Atmosphäre in Schülerlaboren wird als Chance gesehen, die Berührungängste der Schüler mit der Wissenschaft zu verringern (Anton & Hock 2003). Dies scheint auch zu gelingen. Sowohl Oszkan et al. (2008) als auch Zehren (2009) berichten von einer positiven Einschätzung der Merkmale von Laborkursen durch Schüler. Im Rahmen der Lernklima-Forschung wird dem situativen Aspekt der Atmosphäre, also den sozialen Wechselbeziehungen der Lerner, besondere Beachtung geschenkt. Lernklima wird hier als der gemeinsame Erlebensanteil von Unterricht betrachtet. Dieser leitet beim Lerner einen kognitiven Prozess ein, der zu affektiven Urteilen führt (Standop 2002). Einem guten Unterrichtsklima wird zugeschrieben, sich positiv auf Leistungsbereitschaft, Einstellung zum Unterricht, Sozialverhalten und Interessenentwicklung auszuwirken (Gruehn 2000, Meyer 2004). Sowohl im schulischen Unterricht als auch am außerschulischen Lernort Schülerlabor sind also eine positive Atmosphäre sowie ein angenehmes Lernklima anzustreben.

Bisher wurden wichtige Einflussfaktoren für das Lernen am Lernort Schülerlabor vorgestellt. Nachfolgend werden in der Person des Schülers liegende Variablen aufgezeigt, von denen zu erwarten ist, dass diese die Wahrnehmung des Laborkurses beeinflussen bzw. durch den Laborbesuch beeinflusst werden. Da das Experimentieren in den untersuchten Schülerlaborkursen die zentrale Position einnimmt, soll im Anschluss die Bedeutung des Experiments in Wissenschaft, schulischem und außerschulischem Unterricht vorgestellt werden. Abschließend werden einige Modelle zur Erklärung experimenteller Kompetenz vorgestellt. Diese bilden die Grundlage für den im Rahmen der vorliegenden Studie entwickelten Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung FSWEx.

1.2.3 Personenvariablen

Unter dem Begriff der Personenvariablen werden hier sowohl soziodemographische Merkmale wie Geschlecht und Migrationshintergrund als auch die kognitiven sowie affektiven Dispositionen der Schüler subsumiert. Neben dem Geschlecht und dem Migrationshintergrund finden sowohl das Fachwissen, das Interesse, das Selbstkonzept und die Selbstwirksamkeit als auch die epistemologischen Überzeugungen von Schülern Berücksichtigung. Die Auswahl der Variablen ergibt sich aus Erfahrungen vergleichbarer Studien sowie aus den Zielsetzungen des BeLL Bio. Demnach möchte das BeLL Bio Interesse wecken, Lernen ermöglichen, Selbstvertrauen zur Auseinandersetzung mit biologischen Fragestellungen vermitteln und Einblicke in den naturwissenschaftlichen Erkenntnisweg ermöglichen (vgl. Punkt 1.2.1.3.).

1.2.3.1 Soziodemographische Faktoren

Sowohl in den Leistungsvergleichsstudien PISA und TIMSS als auch in der aktuellen Schülerlaborforschung spielt die Untersuchung der Variablen Geschlecht und Migrationshintergrund eine große Rolle. Spezielle Programme in Schulen und außerschulischen Einrichtungen zur Förderung von Mädchen sowie von Schülern mit Migrationshintergrund zeugen von der Aktualität des Themas. Auch in der vorliegenden Untersuchung finden diese Faktoren daher Berücksichtigung. Nachfolgend sollen zentrale Resultate der bisherigen Studien skizziert werden.

1.2.3.1.1 Geschlecht

Geschlechterdifferenzen lassen sich im naturwissenschaftlichen Unterricht über die gesamte Schullaufbahn beobachten. Wie TIMSS 2007 belegt, besitzen männliche Schüler bereits in der Grundschule ein deutlich höheres naturwissenschaftliches Fähigkeitsselbstkonzept als ihre weiblichen Mitschüler (Bos et al. 2008). In der Sekundarstufe I ist für beide Geschlechter ein deutlicher Interessenverfall zu beobachten (Löwe 1987). In der Sekundarstufe II ist zu beobachten, dass weibliche Schüler eher als männliche Schüler dazu neigen, sich für nicht-naturwissenschaftliche Fächer zu entscheiden (Zwick & Renn 2000). Eine Ausnahme stellt hier das Fach Biologie dar. Während Physik und Chemie eine Domäne der Männer sind, wird das Fach Biologie von weiblichen Schülern stärker favorisiert (Zwick & Renn 2000). Als Erklärung für diese geschlechtsspezifischen Präferenzen wäre das Vorliegen von Unterschieden in den naturwissenschaftlichen Fähigkeiten der Geschlechter denkbar. Diese Erklärung schließt sich allerdings bei Betrachtung der PISA 2006-Ergebnisse aus. Hier wurden weder im OECD-Durchschnitt noch in Deutschland signifikante Geschlechterunterschiede in der naturwissenschaftlichen Kompetenz festgestellt. Lediglich in der Spitzengruppe sind die Jungen gegenüber den Mädchen in Deutschland überrepräsentiert (Prenzel

2007). Ursächlich für die bereits zu Beginn der Schullaufbahn existierenden Unterschiede im Fähigkeitsselbstkonzept könnte vielmehr die Prägung weiblicher Schüler von dem kulturell bedingten Stereotyp der maskulinen Naturwissenschaften sein (Brandt 2005). Dieses kann dazu führen, dass Mädchen ihre mathematisch-naturwissenschaftlichen Fähigkeiten unterschätzen und ein negatives Berufsbild entwickeln. Selbst im Falle eines positiven Berufsbildes sehen weibliche Schüler eine naturwissenschaftliche Laufbahn oftmals als unrealistisch an (Biermann et al 2009).

Nicht nur bezüglich der Einschätzung ihrer fachlichen Fähigkeiten, sondern auch in ihrer Herangehensweise an naturwissenschaftliche Fragestellungen unterscheiden sich weibliche und männliche Schüler. Während Mädchen ein höheres Interesse zeigen, sich theoretisch mit biologischen Themen auseinanderzusetzen (z.B. durch Lesen im Biologiebuch), ist das Interesse des praktischen Experimentierens auf Seiten der Jungen höher (vgl. Aufdermauer et al. 2006b). Für Mädchen sind zudem die Zusammenarbeit, der Wissenserwerb und die Selbstständigkeit beim Experimentieren von größerer Bedeutung als für Jungen (Engeln 2006). Das erfolgreiche Experimentieren im Schülerlabor erfordert die Kooperation mit Mitschülern, Eigenaktivität und die Einbeziehung des eigenen theoretischen Vorwissens. Somit wird sowohl die kooperative Präferenz der weiblichen als auch die Vorliebe für praktisches Arbeiten der männlichen Schüler angesprochen. Engeln und Euler (2004) gehen deshalb auch davon aus, dass es Schülerlaboren gelingt, Mädchen und Jungen gleich gut zu fördern. Eine Gender Gap, wie sie gelegentlich im schulischen Unterricht zu beobachten ist, stellen sie beim Arbeiten im Schülerlabor nicht fest. Schülerlabore haben nicht nur das Potential, das Interesse von Mädchen und Jungen gleichermaßen zu fördern (Engeln 2006). Auch werden die Relevanz, die intrinsische Motivation und das Selbstkonzept beim Laborbesuch von Mädchen und Jungen gleich empfunden (Brandt 2005). Euler (2009b) erkennt, dass beide Geschlechter den Laborbesuch mit positiven Gefühlen belegen und ihn als persönlich bedeutsam erleben. Ozkan et al. (2008) gehen sogar davon aus, dass Mädchen generell Laborunterricht positiver wahrnehmen als Jungen. Darüber hinaus gelingt es, durch die Teilnahme an Schülerlaborveranstaltungen bei Mädchen und Jungen gleichermaßen eine – wenn auch nur kurzfristige – Steigerung des fachlichen Selbstkonzeptes zu erreichen (Brandt et al. 2008). Eine weitere Unterstützung der weiblichen Schüler durch Einrichtung monoedukativer Gruppen bei der Laborarbeit konnte Brandt (2005) nicht beobachten. Das praktische Experimentieren im Schülerlabor scheint somit durchaus das Potential zu haben, Geschlechterstereotypen entgegenzuwirken. Auch in dieser Untersuchung ist daher die Berücksichtigung des Einflussfaktors Geschlecht von Interesse.

1.2.3.1.2 Migrationshintergrund

Im Rahmen von PISA 2006 wurden in Deutschland ausgeprägte Kompetenzunterschiede zwischen Schülern mit und ohne Migrationshintergrund festgestellt (Prenzel 2007). Wie TIMSS 2007 bestätigt, finden sich diese bereits auf Ebene der Grundschüler wieder (Bos et al. 2008). Mängel in der Sprachkompetenz werden häufig verantwortlich gemacht für die geringeren schulischen Leistungen von Schülern mit Migrationshintergrund. Ursächlich dafür könnte das häufig geringere Sprachvermögen der Eltern sein (Boos-Nünning 2005). Entsprechend pflegt nur ein relativ geringer Prozentsatz dieser Schüler die Landessprache als Umgangssprache (Bos et al. 2008). Aus dieser Beobachtung ergibt sich die Notwendigkeit, in der Generation des Migrationshintergrundes zu unterscheiden. Wurden nicht beide, sondern nur ein Elternteil im Ausland geboren und wird zu Hause die Sprache des Einwanderungslandes gesprochen, so sind kaum Unterschiede in der naturwissenschaftlichen Kompetenz der Schüler zu beobachten. Weiterhin ist zu berücksichtigen, dass Jugendliche mit Migrationshintergrund auch bei geringerer Leistung ein hohes Interesse für Naturwissenschaften bekunden. Zwischen Jugendlichen mit und ohne Migrationshintergrund konnte für Deutschland in PISA 2006 entsprechend kein Unterschied in der Ausprägung des Interesses beobachtet werden (Prenzel 2007). Auch Brandt (2005) stellt bei seiner Untersuchung im Schülerlabor keine wesentlichen, auf den Migrationshintergrund zurückzuführenden Unterschiede in der affektiven Wahrnehmung des Laborbesuches fest. Umfangreichere Studien zum Einfluss des Migrationshintergrundes auf das Lernen am außerschulischen Lernort Schülerlabor stehen noch aus.

1.2.3.2 Fachwissen

In Bildungsstandards und aktuellen Lehrplänen stellt das Fachwissen neben den prozessbezogenen Kompetenzen eine wichtige Komponente der naturwissenschaftlichen Grundbildung dar (KMK 2005, MSW 2008b). Neben dem rein affektiven Erlebnis möchte das BeLL Bio als außerschulischer Lernort mit enger curricularer Anbindung bei seinen Kursteilnehmern daher auch einen kognitiven Lernerfolg erreichen. Gerade der zu hohe Anteil von Schülern unterer Kompetenzstufen in Deutschland (Rönnebeck et al. 2010) erfordert eine Lernumgebung, die nicht nur *high achiever* anspricht. Schülerlaboren wird dieses Potential zugesprochen. Für sie besteht eine Aufgabe darin, das ideale Maß an Herausforderung zu finden, welches sowohl hochkompetente Schüler aktiviert als auch, zum Beispiel durch stärker angeleitete Experimente, Lernanreize für leistungsschwächere setzt (Engeln & Euler 2004). An Schülerlaborstudien, welche systematisch den Einfluss des Laborbesuchs auf den Wissenserwerb untersuchen, mangelt es allerdings. Freedman (1997) gelangt im Rahmen einer *posttest-only*-Studie im Kontrollgruppendesign zu der Erkenntnis, dass sich Laborunterricht positiv auf das Fachwissen einer sehr heterogenen Schülerschaft auswirkt. Ausreichend abgesichert erscheint diese Aussage aufgrund der Ermangelung eines Pretests allerdings nicht. In der Studie von

Glowinski (2007) wird erhoben, für welche Aspekte der Molekularbiologie die Lernenden einen subjektiv wahrgenommenen Wissenszuwachs angeben. Auch hier erfolgt keine direkte Erfassung der Lernleistung, sondern eine subjektive Einschätzung durch die Schüler. Eine der wenigen systematischen Analysen findet sich bei Scharfenberg (2005). Er vergleicht in seiner Schülerlaborstudie den kurzfristigen und persistenten Wissenserwerb von drei Treatmentgruppen miteinander und mit einer unbehandelten Kontrollgruppe. Während eine Treatmentgruppe einen nicht experimentellen gentechnischen Kurs im Schülerlabor besuchte, wurde in einer anderen Treatmentgruppe entsprechender Unterricht am Lernort Schule durchgeführt. Diese beiden Gruppen wurden mit einer weiteren Treatmentgruppe verglichen, welche im Labor Experimentalunterricht gleichen Inhalts absolvierte. In allen Treatmentgruppen stellt Scharfenberg (2005) fest, dass die Schüler einen persistenten Wissensbestand erwerben, einen Teil der Inhalte aber auch wieder vergessen. Bedeutsame Unterschiede im Lernzuwachs und in der Behaltensleistung der drei Treatmentgruppen beobachtet er nicht.

Als besonders relevant für die Effektivität des außerschulischen Unterrichts wird dessen gelungene Vorbereitung im schulischen Unterricht angesehen (Berck & Graf 2010). Scharfenberg (2005) empfiehlt die enge Anbindung des Laborbesuches insbesondere deshalb, weil das im Unterricht erworbene Vorwissen eine zentrale Bedeutung für den Wissenserwerb in der Experimentalsituation hat. Auf diese Weise kann vermieden werden, dass der Laborunterricht auf eine reine Eventveranstaltung ohne großen Lerneffekt reduziert wird (Sunal et al. 2008). Auch im Sinne des gemäßigten Konstruktivismus erscheint diese Annahme berechtigt. Reinmann und Mandl (2006) betonen in ihrem Ansatz, dass Lernen im Sinne eines Konstruktionsprozesses nur möglich ist, wenn auf bereits vorhandenen Kenntnissen und Fähigkeiten aufgebaut werden kann. Eine positive Wirkung der unterrichtlichen Vorbereitung ist jedoch nicht nur für die kognitive Dimension zu belegen. Auch die positive affektive Wahrnehmung des Laborbesuches kann durch eine ausreichende Vorbereitung des Laborbesuches gesteigert werden (Damerau 2010, Aepfelbach 2011). Schüler mit ausführlicher Vorbereitung messen dem Laborbesuch außerdem eine signifikant höhere persönliche Bedeutung bei (Engeln & Euler 2004). Neben dem Aufbau von Laborkursen auf dem vorangegangenen Unterricht misst Sandoval (2008) auch der unterrichtlichen Nachbereitung Bedeutung bei. Diese hilft den Schülern, ihre Kursergebnisse zu reflektieren (Davies et al. 2008). Trotz dieser einstimmig positiven Effekte entfällt die Vor- und Nachbereitung von Schülerlaborbesuchen in der Praxis fast gänzlich (Engeln et al. 2003, Engeln 2004, Engeln & Euler 2004, Brandt 2005, Sandoval 2008, Euler 2009b). Zudem steht in vielen Laborkursen nicht ausreichend Zeit für die Reflektion der Ergebnisse zur Verfügung (Hofstein & Lunetta 2003). Es ist anzunehmen, dass diese mangelnde Verzahnung des Laborbesuchs mit dem Unterricht zu einer nur kurzfristigen kognitiven und affektiven Wirkung von Schülerlaborbesuchen führen kann. Eine bessere Unterrichtsbindung ließe sich generell dadurch

erreichen, dass Lehrkräfte sich um einen Labortermin bemühen, wenn der Laborkursinhalt zum durchgeführten Unterricht passt. Von Bedeutung ist somit auch, dass die Schülerlabore ihre fachwissenschaftlichen Kursinhalte gut auf die schulischen Unterrichtsgegenstände abstimmen (Brandt 2005).

Das curriculumsnahe Kursangebot des BeLL Bio sichert eine gute Unterrichtsbindung. Darüber hinaus ist die obligatorische versuchsskriptgeleitete Vorbereitung jedes Laborbesuchs integraler Bestandteil des Laborkonzepts. Zwischen der Bereitschaft, sich kognitiv mit einem Lerngegenstand auseinanderzusetzen und dem fachlichen Selbstkonzept sowie der Selbstwirksamkeit wird häufig ein Kausalzusammenhang unterstellt, dessen Richtung nicht klar definiert ist (Moschner & Dickhäuser 2010, Köller & Möller 2010). Die Bedeutung beider Konstrukte wird im folgenden Abschnitt zusammengefasst.

1.2.3.3 Selbstkonzept und Selbstwirksamkeit im Schülerlabor

Unter dem Selbstkonzept wird die Selbsteinschätzung als Resultat externer oder interner Vergleiche verstanden, in denen eine Person andere Personen oder andere Aspekte ihres Selbst als Maßstab setzt (Woolfolk 2008). Bezogen auf das Selbstkonzept im Fach Biologie bedeutet dies, dass Schüler ihre biologischen Fähigkeiten internal, beispielsweise mit ihren eigenen Fähigkeiten in anderen naturwissenschaftlichen oder sprachlichen Fächern abgleichen. Gleichzeitig beziehen sie aber auch ihr soziales Umfeld in diesen Vergleich mit ein und beurteilen ihre eigene Leistungsfähigkeit external im Vergleich zu der ihrer Mitschüler. Resultat dieser Vergleiche ist der häufig berichtete „*Big-Fish-Little-Pond*“-Effekt. Dieser besagt, dass Schüler vergleichbarer Begabung ein höheres Selbstkonzept entwickeln, wenn sie eine leistungsschwächere Lerngruppe besuchen. Ein geringeres Selbstkonzept wird entwickelt, wenn diese Schüler einer leistungsstärkeren Lerngruppe angehören (Marsh 2005, Trautwein & Lüdtke 2005). Die Relevanz des Selbstkonzeptes für Lernen und Leistungsmotivation lässt sich an seiner Position im Erwartungs-Wert-Modell nach Eccles et al. (1983) verdeutlichen. Eccles et al. (1983) postulieren in Ihrem Modell, dass die Leistungsmotivation und die Wahl, welche Aufgaben ein Lerner angeht, in direktem Zusammenhang stehen mit dem subjektiven Aufgabenwert und der Erwartung, die Aufgabe erfolgreich zu bewältigen (Abb. 15).

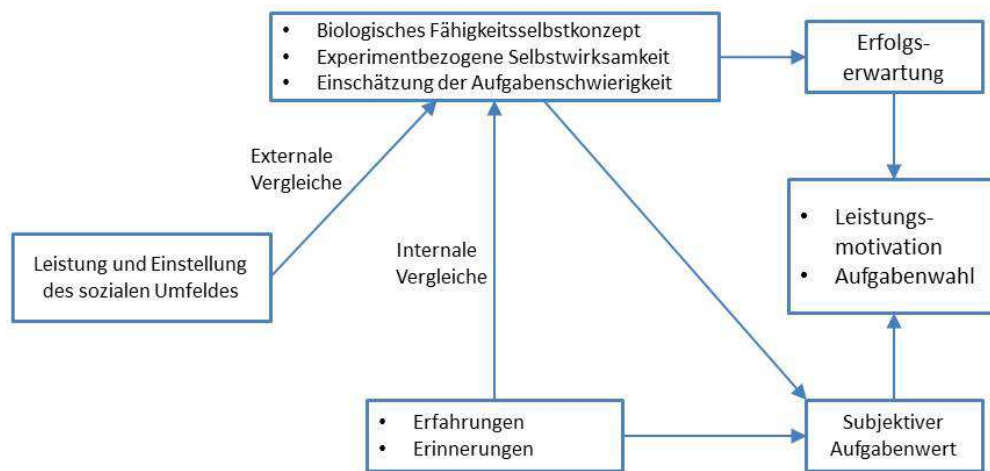


Abbildung 15: Einfluss internaler und externaler Vergleiche auf das Fähigkeitsselbstkonzept und Auswirkungen auf Erfolgserwartung, empfundenen Aufgabenwert und Verhalten (in Anlehnung an das Erwartungs-Wert-Modell von Eccles et al. 1983).

Die Erfolgserwartung einer Person hängt demnach unmittelbar vom Selbstkonzept sowie von der wahrgenommenen Aufgabenschwierigkeit ab. Ebenso beeinflussen diese Faktoren den subjektiven Aufgabenwert, welcher beinhaltet, wie nützlich und wichtig eine Person eine Aufgabe empfindet und wie viel Spaß sie ihr macht (Eccles & Wigfield 1995, Pohlmann et al. 2005). Dem Selbstkonzept kommt somit eine nicht unerhebliche Bedeutung bei der Entwicklung von Leistungsmotivation zu. Sowohl Brandt (2005) als auch Pawek (2009) konnten eine kurzfristige Steigerung des fachlichen Fähigkeitsselbstkonzeptes durch einen einmaligen Schülerlaborbesuch feststellen. Im Falle der Beobachtungen von Brandt im *teutolab*-chemie geht Zehren (2009) allerdings davon aus, dass das hohe Maß an Anleitung während des Experimentierens im *teutolab* eine realistische Selbsteinschätzung der Schüler erschwere.

Die Selbstwirksamkeit ist ein Teilaspekt des globaleren Selbstkonzept-Konstruktes. Der Begriff der Selbstwirksamkeit beschränkt sich auf die selbsteingeschätzte kontextspezifische Kompetenz einer Person, eine bestimmte Aufgabe bewältigen zu können (Köller & Möller 2010). Damit wird der Selbstwirksamkeit einer Person auch zugeschrieben, ein wichtiger Prädiktor für ihr Verhalten zu sein (Bandura 1982). Schätzt eine Person ihre Selbstwirksamkeit hoch ein, so setzt sie sich dementsprechend höhere Ziele, was sich positiv auf ihre Anstrengungsbereitschaft und damit auf ihre Motivation auswirkt (Woolfolk 2008). Es ist also schlüssig, der Selbstwirksamkeit im oben genannten Modell (Abb. 15) eine gleichrangige Position zuzuordnen wie dem Selbstkonzept. In Deutschland bescheinigen sich die Jungen eine höhere Selbstwirksamkeit als die Mädchen. Entsprechend der vorherigen Annahme findet sich dieser Geschlechterunterschied auch in deren Motivation wieder. Sowohl die extrinsische als auch die intrinsische Motivation sind bei den Mädchen in Deutschland geringer ausgeprägt als bei den Jungen (Prenzel 2007).

Selbstwirksamkeit ist nicht statisch. Erlebt sich eine Person in einer Situation als kompetent oder wird ihr von ihrer Umgebung positive Rückmeldung gegeben, so wirkt sich dies positiv auf die Einschätzung ihrer Selbstwirksamkeit aus (Woolfolk 2008). So ist anzunehmen, dass auch das Kompetenzerleben von Schülern beim Experimentieren im Schülerlabor zu einer gesteigerten Selbsteinschätzung führen kann. Dies setzt Experimente voraus, die es Schülern erlauben, ihr Vorwissen – sei es theoretischer oder praktischer Art – einzubringen. Konstruktive und positive Rückmeldungen von Seiten der Kursbetreuer können diese Entwicklung unterstützen. Das Empfinden von Kompetenz wird zudem als notwendige Voraussetzung für die Entstehung intrinsischer Motivation angesehen (Ryan und Deci 2000). Wie Krapp und Ryan (2002) feststellen, muss eine hohe Selbstwirksamkeit allerdings nicht notwendigerweise mit einem hohen Interesse einhergehen, da die Selbstwirksamkeit zwar eine notwendige, aber keine hinreichende Bedingung der Interessenentwicklung darstellt. Auch Engeln (2004) beobachtet in ihrer Schülerlaborstudie entsprechend nur mäßige Korrelationen zwischen dem Fähigkeitsselbstkonzept und der emotionalen, epistemischen und wertebezogenen Komponente des aktuellen Interesses.

1.2.3.4 Interessengenese im Schülerlabor

Ein grundlegendes Interesse von Jugendlichen an den Naturwissenschaften wird als Voraussetzung für eine erfolgreiche Teilhabe am gesellschaftlichen Leben gesehen (Prenzel 2007). Wie Löwe (1987) feststellt, nimmt das Interesse von Schülern an Themen der Biologie in der Sekundarstufe deutlich ab. Zudem ist zu beobachten, dass ein nicht geringer Anteil hochkompetenter Jugendlicher geringes Interesse an Naturwissenschaften zeigt (Prenzel 2007). Gerade weil das Interessenprofil zu den besten Prädiktoren für die spätere Berufswahl zählt (Ludwig 2003), ist eine gezielte Interessenförderung von besonderer Bedeutung. Wie Engeln (2004) feststellt, haben Schülerlabore das Potential, Interesse für Naturwissenschaften und Technik zu erzeugen. Somit ist eine Berücksichtigung des Interesses auch in dieser Studie erforderlich.

Im Zentrum moderner Interessentheorien steht die Person-Gegenstands-Beziehung. Diese führt Interesse auf den ständigen Austausch einer Person mit ihrer sozialen und gegenständlichen Umwelt zurück. Eine interessenthematische Beziehung zeichnet sich demnach durch das Erleben von positiven emotionalen Zuständen (emotionales Interesse), durch eine subjektiv hohe Wertschätzung (wertebezogenes Interesse) sowie durch das Bestreben zunehmender Wissensausdifferenzierung (epistemisches Interesse) aus. Für das Lernen ist insbesondere das epistemische Interesse von Bedeutung, da eine epistemisch interessierte Person sich mit dem Gegenstand des Interesses identifiziert und mehr über ihn erfahren möchte. Zu unterscheiden ist weiterhin zwischen persönlichem (=individuellem) und situationalem Interesse. Während persönliches Interesse ein

dispositionales Persönlichkeitsmerkmal ist, rührt situationales Interesse aus den Anregungsbedingungen einer Situation, zum Beispiel einer außerschulischen Lernumgebung, her. Das aktuell vorhandene Interesse ergibt sich somit aus der Interessantheit der Situation sowie der Aktualisierung des persönlichen Interesses. Identifiziert sich die Person mit dem Lerngegenstand, kann ihr dispositionales Interesse langfristig stabilisiert werden (Krapp 2010, Abb. 16). Mitchel (1993) differenziert in Bezug auf die Entwicklung situationalen Interesses in eine *catch*- und in eine *hold*-Komponente. Eine Lernumgebung fängt zunächst kurzfristig die Aufmerksamkeit des Lernalers ein. Dies gelingt ihr durch *catch*-Komponenten wie Überraschung, Diskrepanzerleben oder die Einbettung von Lerninhalten in interessante Kontexte. Eine nachhaltige Stabilisierung des situationalen Interesses im Sinne einer *hold*-Komponente wird erst dadurch erreicht, dass der Lerngegenstand als etwas Sinnvolles erlebt wird (Krapp 1998, Abb. 16).

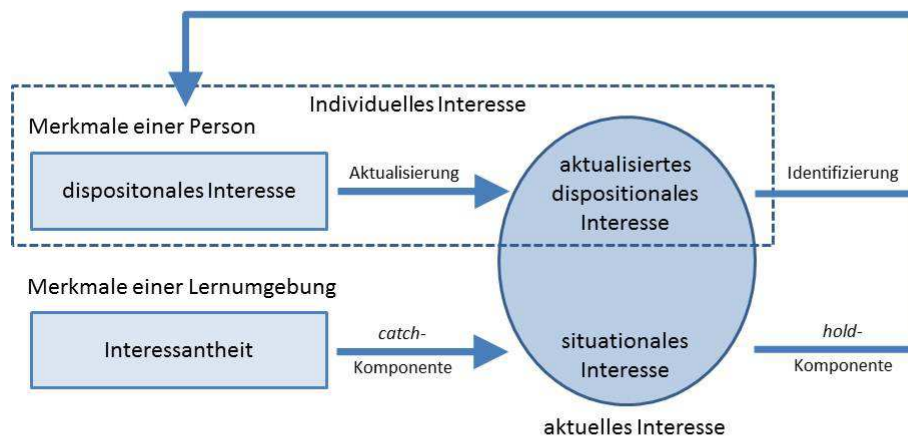


Abbildung 16: Konzepte von Interesse (verändert nach Engel 2004)

Ein möglicher Einfluss des Interesses wird darin gesehen, dass es zu einer stärkeren kognitiven Aktivierung und damit zu einer höheren Behaltensleistung führt (Krapp 2010). Im Fach Biologie finden Schiefele et al. (1993) allerdings nur geringe Korrelationen zwischen Fachinteresse und Lernerfolg. Bezogen auf das Experimentieren wird davon ausgegangen, dass sich Interesse insbesondere bei der Planung und Durchführung von Experimenten entwickelt, während naturwissenschaftliche Kompetenz vor allem dann entsteht, wenn in hohem Umfang Schlussfolgerungen aus Experimenten ermöglicht werden (Seidel et al. 2007). Da in vielen Schülerlaborkursen ein Schwerpunkt auf der Durchführung und Auswertung von Experimenten liegt, sollten diese somit das Potential besitzen, sowohl Interesse hervorzurufen als auch zum Lernerfolg beizutragen. Allerdings betonen Guderian et al. (2006a), dass sowohl ein einmaliger als auch ein wiederholter Schülerlaborbesuch im Sinne einer *catch*-Komponente nur zu einer kurzfristigen Steigerung des Interesses führen können. Im Anschluss an den Besuch liegt die Verantwortung in den Händen der Lehrkraft, daraus ein langfristiges Interesse aufzubauen (Guderian et al. 2006a). Entsprechend stellt Engel (2004) in ihrer Schülerlaborstudie fest, dass das emotionale und

epistemische Interesse innerhalb von zwölf Wochen nach einem einmaligen Laborbesuch wieder absinken. Auch Pawek (2009) beobachtet, dass unmittelbar nach einem Laborbesuch der Einfluss der Lernumgebung auf das aktuelle Interesse vergleichbar stark ist wie der Einfluss des dispositionalen Interesses. Sechs bis acht Wochen nach dem Laborbesuch sinkt allerdings der Einfluss der Lernumgebung und das bereits vor dem Laborkurs bestehende dispositionale Interesse dominiert wieder. Der Einfluss eines Laborbesuches auf das Berufsinteresse wird von Brandt (2005) in seiner Studie im Bielefelder *teutolab*-chemie berücksichtigt. Hier stellt er fest, dass die Schüler sich nach einem Laborbesuch eher vorstellen können, in einem Beruf zu arbeiten, der mit Chemie zu tun hat.

1.2.3.5 Epistemologische Überzeugungen

„Epistemologische Überzeugungen sind Ansichten über die Praxis naturwissenschaftlichen Arbeitens, den epistemologischen Status naturwissenschaftlichen Wissens, die Strukturierung und Klassifikation von Wissen und die persönliche Bedeutung naturwissenschaftlicher Inhalte.“ (Priemer 2006). Der Begriff wird häufig synonym mit dem des Wissenschaftsverständnisses verwendet. Da in Schülerlaboren das praktische Experimentieren im Vordergrund steht, wird in der vorliegenden Arbeit unter Wissenschaftsverständnis insbesondere die angemessene Einschätzung der Rolle von Experimenten für den naturwissenschaftlichen Erkenntnisprozess verstanden (Urhahne und Hopf 2004). Häufig haben Schüler nur sehr stereotype Vorstellungen von der Natur der Naturwissenschaften (Wehlen 2007), was insbesondere für diejenigen gilt, die aufgrund ihres sozialen oder schulischen Umfeldes wenig Einblick in entsprechende Berufsfelder erhalten (Biermann et al. 2009). Beispielsweise wird das Labor häufig als wichtiger Arbeitsplatz eines Wissenschaftlers gesehen, nicht aber sein Büro. Die Tätigkeit des Experimentierens wird bei jüngeren Schülern häufig als planloses Ausprobieren aufgefasst, während für ältere Schüler das Sammeln von Daten im Vordergrund steht (Höttecke 2001). Diese reduzierten Sichtweisen stimmen insofern bedenklich, da reifere epistemologische Überzeugungen mit einem höheren fachspezifischen Selbstkonzept sowie mit einer höheren Leistungsmotivation und tiefergehenden Lernstrategien einhergehen (Urhahne & Hopf 2004, Urhahne 2006). Somit ist anzunehmen, dass sich die Qualität des dispositionalen Wissenschaftsverständnisses auch auf die Wahrnehmung der Lernumgebung Labor sowie im Experimentierverhalten auswirkt.

Eine mögliche Ursache für die Mängel im Wissenschaftsverständnis der Schüler ist in ihrem möglicherweise von der Lehrkraft deutlich abweichenden Wissenschaftsbild zu sehen. Dieses kann zu Missverständnissen im Schüler-Lehrer-Dialog und damit zu mangelndem Lernerfolg führen (Leach 1998). Auch verwundert es nicht, wenn die Bedeutung des Experimentierens für den wissenschaftlichen Erkenntnisweg nicht erkannt wird, da dem eigenständigen Experimentieren in der

Schule nur wenig Bedeutung zukommt. Dabei kann gerade dieses praktische Arbeiten Schülern einen Einblick in den wissenschaftlichen Forschungsprozess vermitteln (Leach 1998). Aus epistemologischer Sicht kommt daher dem praktischen, von echten Wissenschaftlern betreuten, Arbeiten in Schülerlaboren eine wichtige Rolle zu. Dieser originale Kontakt eröffnet Einblicke in den wissenschaftlichen Erkenntnisprozess, bietet die Chance zum Abbau stereotyper Vorstellungen und trägt damit zu einem realistischeren Verständnis der Arbeit eines Wissenschaftlers bei. Insbesondere durch das kontinuierliche Einbringen epistemologischer Aspekte in den Unterricht könnte zur Entwicklung eines Wissenschaftsverständnisses beigetragen werden (Günther et al. 2003). Regelmäßig in den naturwissenschaftlichen Unterricht integrierte Unterrichtsgänge zu Schülerlaboren scheinen daher zur Förderung des Wissenschaftsverständnisses beizutragen. Dies sollte sowohl für weibliche als auch männliche Schüler gelten, denn Geschlechtsunterschiede sind in den epistemologischen Überzeugungen nicht zu beobachten (Urhahne & Hopf 2004).

1.2.4 Experimentieren

Als eine zentrale Methode der Erkenntnisgewinnung kommt dem Experiment in den Naturwissenschaften eine besondere Rolle zu (Lehrer et al. 2001, Köhler 2004, Gentürk 2009). Wie die folgenden Abschnitte zeigen, wird dem Experimentieren auch in den Bildungsstandards (KMK 2005), in den Lehrplänen und im naturwissenschaftlichen Unterricht entsprechende Bedeutung beigemessen. Die mitunter abweichende Nutzung des Experimentbegriffs in Wissenschaftstheorie und Fachdidaktik machen eine differenziertere Betrachtung erforderlich. In Anlehnung an Gropengießer und Kattmann (2008) wird der Begriff Experiment in der vorliegenden Arbeit synonym mit der Bezeichnung Versuch verwendet.

1.2.4.1 Das Experiment in der Wissenschaftstheorie

Eine Grundüberzeugung der Naturwissenschaften ist es, dass Naturphänomene auf natürlichen erklärbaren Ursachen beruhen. Die Klärung dieser Ursachen geschieht im Prozess der naturwissenschaftlichen Erkenntnisgewinnung, welcher häufig durch hypothetisch-deduktives Denken geprägt ist. In diesem wird zunächst – meist theoriegeleitet – eine Hypothese aufgestellt, welche anschließend in Form eines Experiments oder einer Beobachtung getestet wird (Campbell 2000, Graßhoff et al. 2000). In einem naturwissenschaftlichen Experiment wird nach einem von Hypothesen geleiteten Versuchsplan eine Situation zu Beobachtungszwecken hergestellt. Darin soll untersucht werden, ob ein zu beobachtender bzw. zu messender Zusammenhang mit einer hypothetischen Einflussgröße besteht und wie dieser aussieht (Gropengießer & Kattmann 2008). Als vehementer Vertreter dieser deduktiven Sichtweise bemühte Karl Popper sich, die Idee des empirisch-induktiven Vorgehens zu widerlegen (Popper & Miller 1983). Im Gegensatz zur Deduktion

wird bei letzterem Vorgehen vom Speziellen auf das Allgemeine geschlossen. Aus einer gewissen Anzahl an Beobachtungen oder Experimenten kann dem induktiven Ansatz zufolge eine Theorie abgeleitet werden (Backhaus & Braun 2009). Ein Blick in die Bildungsstandards eröffnet, dass hier eher der hypothetisch-deduktive Charakter des Experiments betont wird.

1.2.4.2 Das Experiment in Bildungsstandards und Lehrplänen des Unterrichtsfachs Biologie

Wie bereits angemerkt, führte das mangelhafte Abschneiden deutscher Schüler bei PISA 2000 und TIMSS auf bildungspolitischer Ebene zu deutlichen Reaktionen. Zu diesen zählt auch die vom Bundesministerium für Bildung und Forschung (BMBF) in Auftrag gegebene Expertise zur Entwicklung nationaler Bildungsstandards, auch als Klieme-Gutachten bekannt. Bereits in diesem, auf fachlicher Ebene wenig konkreten Gutachten wird die Umsetzung von Maßnahmen im Unterricht empfohlen, die zum Erwerb prozeduralen Handlungswissens beitragen (Klieme et al. 2007). In den auf Grundlage des Gutachtens erlassenen Bildungsstandards für den mittleren Schulabschluss im Fach Biologie finden sich diese Empfehlungen in den vier Kompetenzbereichen wieder. Die Schüler sollen demnach bis zum mittleren Schulabschluss über Kompetenzen in den Bereichen Fachwissen, Kommunikation, Erkenntnisgewinnung und Bewertung verfügen. Neben dem Beobachten, Vergleichen und Anwenden von Modellen beinhaltet der Kompetenzbereich Erkenntnisgewinnung explizit die Kompetenzen des Anwendens von Arbeitstechniken sowie des Experimentierens (KMK 2005). Diese Kompetenzen sollen es Schülern erlauben, den naturwissenschaftlichen Erkenntnisgewinn nachzuvollziehen und zu verstehen, was naturwissenschaftliche Methodik und Aussagen charakterisiert (Mayer 2007). Experimentieren wird in den Bildungsstandards auch näher als hypothesengeleitetes Experimentieren bezeichnet, bei dem die Schüler in drei Schritten vorgehen: Zunächst formulieren sie aus einem Problem heraus eine Fragestellung und stellen eine Hypothese auf. Im zweiten Schritt planen sie eine Beobachtung, einen Vergleich oder ein Experiment und führen dies durch. Zuletzt sollen die experimentell gewonnenen Daten ausgewertet und hinsichtlich der im ersten Schritt aufgestellten Hypothesen interpretiert werden (KMK 2005).

Die Implementierung der Bildungsstandards durch die Bundesländer führte dazu, dass dem Experiment auch in den (Kern-)Lehrplänen zentrale Bedeutung zukommt. Wenn auch bisher für die Grundschule keine Bildungsstandards im Fach Sachunterricht existieren, so wird hier doch erwartet, dass die Schüler lernen zu experimentieren. Am Ende der Schuleingangsphase sollen sie die Eigenschaften von Wasser, Luft, Wärme und Kälte, Licht und Schatten in Experimenten entdecken können. Mit Abschluss der vierten Klasse sollen sie bereits befähigt sein, erste Versuche zu Themen wie Licht, Feuer, Wasser, Luft oder Schall zu planen, durchzuführen und auszuwerten (MSW NRW 2008a). Im Kernlehrplan der Hauptschule ist sehr ausdifferenziert niedergelegt, wie die

experimentelle Kompetenz am Ende der ersten (in Biologie in der Regel nach etwa der Hälfte, in Chemie und Physik nach etwa einem Drittel der bis Ende der Klasse zehn vorgesehen Unterrichtszeit) und der zweiten Regressionsstufe (mit Abschluss der Klasse zehn) in den Naturwissenschaften aussehen soll (MSW NRW 2011a). Ähnliche Darstellungen finden sich auch in den Kernlehrplänen für die Sekundarstufe I der Realschule (MSW NRW 2011c) und des Gymnasiums im Fach Biologie. Mit Abschluss der Sekundarstufe I sollen die Gymnasiasten im Kompetenzbereich Erkenntnisgewinnung beispielsweise Fragestellungen erkennen können, die mit Hilfe biologischer Untersuchungen zu beantworten sind. Sie sollen Hypothesen aufstellen, geeignete Experimente planen, durchführen, auswerten und protokollieren können (MSW NRW 2008b). Im Rahmen der gymnasialen Oberstufe sollen die Methoden der naturwissenschaftlichen Erkenntnisgewinnung vertieft werden. Insbesondere der praktischen, methodenorientierten Arbeit wird in diesem Zusammenhang ein hoher Stellenwert beigemessen. Weiterhin wird im Lehrplan der Oberstufe dem praktischen Arbeiten das Potential zugeschrieben, sozialintegratives Verhalten und Teamfähigkeit zu schulen (MSWWF NRW 1999). Ob der zentralen Stellung des Experiments in den Lehrplänen auch in der Praxis des naturwissenschaftlichen Unterrichts ausreichend Beachtung geschenkt wird, wird der nächste Abschnitt zeigen.

1.2.4.3 Bedeutung des Experiments im Unterricht

Das Experiment nimmt nicht nur in den Lehrplänen, sondern auch in der fachdidaktischen Literatur und in Schulbüchern eine zentrale Position ein (Euler 2001). Praktisch experimentiert wird dagegen nur relativ selten (Barth & Pfeiffer 2009), so dass Schüler häufig wenig Experimentiererfahrung haben (Engeln 2004). Gerade für die Entwicklung von naturwissenschaftlichen Kompetenzen und Interessen scheint es aber förderlich zu sein, wenn die Gelegenheit zum eigenständigen Experimentieren einen ähnlichen Unterrichtsanteil erhält, wie das Schlussfolgern, das Entwickeln eigener Ideen und das Übertragen von Konzepten auf den Alltag (Prenzel 2007). Darüber hinaus ist zu beobachten, dass die im naturwissenschaftlichen Unterricht erlernten praktischen Abläufe oft wenig gemein haben mit den in echten Laboren gefragten Fähigkeiten (Osborne 1998). Die wenigen Experimente sind außerdem häufig schlecht in den Unterricht integriert, so dass ihr Potential zur Aktivierung der Schüler kaum genutzt wird (Bell & Hackl 2006). Pfangert-Becker (2010) betont die Notwendigkeit, das Unterrichts-Experiment einzubetten in den Prozess der naturwissenschaftlichen Erkenntnisgewinnung. Im Rahmen hypothetisch-deduktiven oder empirisch-induktiven Vorgehens lernt der Schüler das Experiment als Instrument naturwissenschaftlichen Arbeitens kennen. Darüber hinaus kann das Experiment aber auch als Objekt und Medium des Lernens aufgefasst werden. Als Objekt des Lernens repräsentiert das Experiment selbst den Lerngegenstand. Gleichmaßen ist das Experiment aber auch Medium des Lehrens, da es selbst den Lernprozess in Gang setzt und aufrecht

erhält. Wird das Experiment problem- und handlungsorientiert eingesetzt, so bietet es den Schülern die Möglichkeit, aktiv und situiert deklaratives, prozedurales und konditionales Wissen zu erwerben (Pfanger-Becker 2010). Die vielfältige Bedeutung des Unterrichtsexperiments wird in Abbildung 17 zusammengefasst.



Abbildung 17: Das Experiment im naturwissenschaftlichen Lehr-/Lernprozess (nach Pfanger-Becker 2010).

Verschiedene Formen des Experimentierens spielen im Biologieunterricht eine Rolle. Während qualitative Experimente sich in allen Jahrgangsstufen finden, sind quantitative Experimente gewöhnlich, bedingt durch das erforderliche Vorwissen und den höheren messtechnischen Aufwand, auf mittlere oder höhere Jahrgangsstufen beschränkt (Killermann et al. 2009). Differenziert man das Unterrichtsexperiment nach seiner äußeren Form, so können Experimente in Lehrerdemonstrationen (mit oder ohne Schülerbeteiligung) und Schülerversuche unterteilt werden (Pfanger-Becker 2010). Während im Sachunterricht der Grundschule noch mehr Schüler- als Lehrerversuche durchgeführt werden (Bäumel-Roßnagl 1981), kehrt sich dieses Bild in weiterführenden Schulen um. Hier dominieren Demonstrationsexperimente durch die Lehrkräfte, während Schülerexperimente und Gelegenheiten zum selbstständigen Forschen kaum vorzufinden sind (Prenzel 2007). Während Lehrkräfte im Physikunterricht eine recht hohe Bereitschaft zur Durchführung von Schülerexperimenten signalisieren, kommen diese im Biologieunterricht deutlich seltener vor (Barth, Pfeiffer 2009). Dies sollte zu denken geben, da Schülerexperimenten zugesprochen wird, zur Selbstständigkeit und manuellen Geschicklichkeit beizutragen, die Motivation zu fördern und ein Gefühl für fachliche Arbeitsweisen zu vermitteln (Backhaus & Braun 2009). Oft werden Material- und Zeitmangel oder Sicherheitsaspekte als Gründe für Demonstrationsversuche angeführt (Killermann et al. 2009). Neben Kurzzeitexperimenten, die in einer Einzel- oder Doppelstunde durchführbar sind, werden im Biologieunterricht auch langzeitige, zum Teil über Tage oder Wochen andauernde

Versuche durchgeführt. Bei diesen muss die Lehrkraft ihr besonderes Augenmerk darauf richten, die Aufmerksamkeit der Schüler aufrecht zu erhalten (Killermann et al. 2009).

Als Modellexperimente werden solche Versuche bezeichnet, die nicht am Original vorgenommen werden. Einen besonders hohen Abstraktionsgrad weisen Gedankenexperimente auf. Bei diesen wird ein Versuch rein theoretisch verbal oder schriftlich durchdacht (Gropengießer & Kattmann 2008). Zunehmend finden heute auch immer mehr virtuelle Experimente in Form von Computersimulationen Anwendung (Killermann et al. 2009). Im Unterricht kann das Experiment verschiedene didaktische Funktionen erfüllen. Als Einführungsversuch dient es als Einstieg in eine neue Fragestellung und wird von der Lehrkraft eingesetzt, um auf bestimmte Phänomene aufmerksam zu machen. Das bestätigende Experiment wird hingegen eingesetzt um Sachverhalte, die den Lernenden bereits bekannt sind, zu bestätigen, zu wiederholen und zu vertiefen. In klärenden beziehungsweise entdeckenden Experimenten untersuchen Schüler etwas für sie persönlich Neues, auch wenn das Ergebnis in Forscherkreisen bereits bekannt ist (Gropengießer & Kattmann 2008). Beim Durchführen historischer Experimente können sie beispielsweise Entdeckungen nachvollziehen, die in der Vergangenheit gemacht wurden (Wellington 1998). Generell kann diese Form des Experiments der Lösung aufgeworfener Fragen oder Probleme dienen. Ihr kommt damit die größte Nähe zum Forschungsexperiment zu (Killermann et al. 2009; Tab. 3).

Tabelle 3: Einteilung der Experimente nach verschiedenen Kriterien (in Anlehnung an Killermann et al. 2009)

Kriterium	Form der Experimente		
Genauigkeit der Ergebnisse	qualitativ		quantitativ
Ausführenden Person	Lehrkraft	Demonstration (Lehrkraft + Schüler)	Schüler
Zeitbedarf	gering		hoch
Abstraktionsebene	mit realen Objekten	mit realen (Funktions-) Modellen	mit virtuellen Modellen/Gedankenexperimente
Didaktische Funktion	einführend	klärend/entdeckend	bestätigend

Mit dem Experimentieren kann die Lehrkraft im Unterricht verschiedene Ziele verfolgen. Während das praktische Arbeiten im naturwissenschaftlichen Unterricht unterer Jahrgangsstufen insbesondere Primärerfahrungen mit naturwissenschaftlichen Phänomenen ermöglichen soll, hat das Experimentieren in weiterführenden Jahrgängen, insbesondere in der Oberstufe, auch

wissenschaftspropädeutische Funktion (MSWWF NRW 1999). Bei Schülern im Chemieunterricht der Sekundarstufe I stellt Pfangert-Becker (2010) zudem fest, dass im Experimentalunterricht gelegentlich Bedenken vor dem Misslingen eines Versuches einschließlich der daraus resultierenden Konsequenzen (z.B. Brände, Explosionen, Verätzungen) geäußert werden. Er sieht es daher als ein primäres Ziel an, diese Ängste abzubauen und das Interesse zu fördern. Darüber hinaus artikuliert er das Ziel, ein Grundverständnis zu erreichen, welches dem Lernenden eine naturwissenschaftlich orientierte Auseinandersetzung mit seiner Lebenswelt ermöglicht (Pfangert-Becker 2010). Backhaus und Braun (2009) fassen die Ziele des Experimentierens folgendermaßen zusammen: Lernende sollen beim Experimentieren

- Theorie und Praxis miteinander verbinden,
- die Fähigkeit erwerben, mit experimentellen Versuchsgescherten umzugehen,
- Methoden wissenschaftlichen Denkens und Handelns kennenlernen und anwenden,
- ein Gefühl für physikalische [biologische] Phänomene bekommen,
- die Möglichkeit haben, ihre eigene Meinung zu artikulieren und die Vorstellungen anderer zu erleben und zu bewerten (Backhaus und Braun 2009).

Die Ansichten über die tatsächliche Wirksamkeit des Experimentierens im Unterricht sind sehr heterogen (Engeln & Euler 2004, Euler 2009b). Aufdermauer & Hesse (2006a) gehen davon aus, dass fachgemäße Arbeitsweisen im Biologieunterricht zur Schulung des Wahrnehmungs-, des Denk- und Urteilsvermögens sowie des Sozialverhaltens beitragen und sich positiv auf die Motivation, das Interesse und die Selbstständigkeit von Schülern auswirken. Einen Hinweis auf einen möglichen Zusammenhang von Experimentierhäufigkeit und Interesse liefert Prenzel (2007). Im Kontext der PISA 2006-Ergebnisse berichtet er von einem tendenziellen, positiven Zusammenhang zwischen der Experimentierhäufigkeit im Unterricht, die Jugendliche angeben, und ihrem Interesse. Generell wird dem Experimentieren insbesondere dann eine positive Wirkung beigemessen, wenn bei der Versuchsdurchführung die selbstständige Erschließung naturwissenschaftlicher Erkenntnisse ermöglicht wird. Leisen (2006) sieht darin die Chance, Schüler auch kognitiv zu aktivieren. Eine deutlich kritischere Haltung nimmt Osborne (1998) ein. Er glaubt, dass die Bedeutung des praktischen Arbeitens im naturwissenschaftlichen Unterricht überschätzt wird und geht davon aus, dass Schüler eher durch häufigeres wissenschaftliches Lesen, Schreiben und Diskutieren profitieren können. Dem Erlernen eines breiten Spektrums an Labormethoden misst er nur Bedeutung bei für Schüler, die später eine wissenschaftliche Karriere einschlagen wollen. Festzuhalten bleibt, dass Experimentieren allein noch kein Garant für Motivation und Lernerfolg ist (Engeln & Euler 2004). Eine nachhaltige Wirkung im Sinne der oben genannten Ziele des Experimentierens sollte nur dann

erwartet werden, wenn die Experimente in geeigneter Weise in den Unterricht eingebunden werden (Euler 2001).

1.2.4.4 Probleme beim Experimentieren im Unterricht

Trotz des hohen Potentials, welches im schulischen Experimentieren gesehen wird, wird im Alltag des Biologieunterrichts relativ selten experimentiert (Scharfenberg 2005). Gründe dafür sind sowohl objektiver als auch subjektiver Natur (Berck & Graf 2010). Als zentrale Probleme objektiver Art werden von Lehrern häufig ungeeignete Fachräume, mangelnde Ausstattung, zu hohe Kosten und ein zu hoher Zeitaufwand genannt (Bäumel-Roßnagl 1981, Scharfenberg 2005). Gerade umfangreichere Versuche – wie sie in der Oberstufe häufig von Interesse wären – sind unter diesen Bedingungen kaum realisierbar. Steht den Leistungskursen immerhin ausreichend Zeit zur Verfügung, um theoretisch Experimente durchführen zu können, beklagen Grundkurse häufig Zeitnot (Wein 2009). Weitere limitierende Faktoren könnten sowohl in der mangelnden Praxiserfahrung der Schüler (Sunal et al. 2008) als auch in der Experimentierfähigkeit der Lehrperson liegen (Scharfenberg 2005). Kommt diese zusammen mit Mängeln in der fachlichen und/oder didaktischen Ausbildung (vgl. acatech 2011), so ist zu erwarten, dass diese Lehrperson vom Experimentieren im Unterricht eher Abstand nimmt. Zudem wird eine sich subjektiv unsicher einschätzende Lehrperson ihren Zeitaufwand für die Planung, Durchführung und Nachbereitung von Experimenten höher einschätzen. Es ist weiterhin anzunehmen, dass ihre subjektive empfundene Experimentierfähigkeit auch auf die Experimentiersicherheit ihrer Schüler abfärbt. Auch die rechtlichen Vorschriften und Regeln zur Sicherheit beim Experimentieren in der Schule können ein limitierender Faktor sein (Scharfenberg 2005). Resümierend lässt sich festhalten, dass das Potential des Experimentierens sich unter schulischen Bedingungen kaum voll entfalten kann. Insbesondere die mit vielen Experimenten einhergehenden räumlichen und materiellen Anforderungen können häufig nur von speziell dafür ausgerüsteten Schülerlaboren erfüllt werden (Brandt 2005).

1.2.4.5 Modelle experimenteller Kompetenz

Im vorangegangenen Abschnitt wurde die geringe Experimentierfähigkeit von Schülern und Lehrkräften als eine mögliche Ursache für die geringe Experimentierhäufigkeit im Unterricht genannt. Doch was ist überhaupt Experimentierkompetenz? Eine nähere Definition dieses Begriffs macht es zunächst notwendig, sich mit dem Kompetenzbegriff auseinanderzusetzen. Häufig wird dessen Umschreibung nach Weinert (2001) zitiert. Demnach sind Kompetenzen die kognitiven Fähigkeiten und Fertigkeiten, die den Individuen entweder verfügbar oder von ihnen erlernbar sind, bestimmte Probleme zu lösen. Dazu gehören auch die damit verbundenen motivationalen, volitionalen und

sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können (Weinert 2001).

Neben einer kognitiven Komponente und motivationalen Orientierungen erfasst diese Definition auch insbesondere die zur Bewältigung von alltäglichen Problemsituationen relevanten anwendungsbezogenen Fertigkeiten. Diese Handlungsdimension wird in den Bildungsstandards des Faches Biologie (KMK 2005) von den drei Kompetenzbereichen Erkenntnisgewinnung, Kommunikation und Bewertung repräsentiert. Hier wird das Experimentieren explizit als Teilkompetenz des Bereiches Erkenntnisgewinnung erfasst. Der Modellierung naturwissenschaftlicher Kompetenzen wird insbesondere seit der Einführung der nationalen Bildungsstandards zunehmend Bedeutung in der fachdidaktischen Forschung beigemessen (Schecker & Parchmann 2006). Kompetenzmodelle bemühen sich auf der Basis pädagogisch-psychologischer und fachdidaktischer Forschung um die nähere Ausdifferenzierung und Beschreibung von Kompetenzen. Klieme et al. (2007) differenzieren zwischen Komponentenmodellen und Stufenmodellen. Während Erstere das Gefüge der Anforderungen beschreiben, welches von Schülern erwartet wird, differenzieren Stufenmodelle (zusätzlich) zwischen Kompetenzgraden. Eine weitere Differenzierung in Kompetenzstruktur- und Entwicklungsmodelle erscheint sinnvoll. Kompetenzentwicklungsmodelle beziehen im Gegensatz zu Strukturmodellen die zeitliche Entwicklung von Kompetenzen (z.B. über Alters- oder Schulstufen hinweg) in die Modellstruktur ein.

1.2.4.5.1 Scientific Discovery as Dual Search (SDDS)-Modell (Klahr & Dunbar 1988)

Klahr und Dunbar (1988) beschreiben das kognitionspsychologische SDDS-Modell „[...] as a general model of scientific reasoning, that can be applied to any context in which hypotheses are proposed and data are collected.“ (Klahr und Dunbar 1988). Das Modell erhebt den Anspruch, allgemeingültige Aussagen über die Vorgänge des naturwissenschaftlichen Erkenntnisprozesses ableiten zu können. Durch die Komponenten des Modells wird der Erkenntnisprozess von der anfänglichen Hypothesenbildung, über die folgende experimentelle Überprüfung der Hypothesen, bis zur finalen Analyse der experimentell generierten Evidenzen erfasst (Hammann 2007). Grundannahme des Modells ist die Betrachtung des Erkenntnisprozesses als Problemlösevorgang. Ein Problem besteht demnach aus einem Ausgangszustand (dem ungelösten Problem), einem Zielzustand (dem gelösten Problem) und aus einem Set von Operatoren, Beschränkungen und Zwischenzuständen. Die Gesamtheit dieser Bestandteile bezeichnet Klahr (2000) als Problemraum (*problem space*). Den Problemlösevorgang sieht er als Suche eines Pfades durch den Problemraum, welcher den Ausgangszustand in den Zielzustand überführt. Die Gewinnung naturwissenschaftlicher Erkenntnisse erfordert nach Klahr (2000) die Suche in zwei Problemräumen, dem Hypothesen- (*hypothethis space*)

und dem Experiment-Raum (*experiment space*). Ausgangszustand des Hypothesenraumes bildet das spezifische Vorwissen über ein zu klärendes Phänomen. Durch *Suche im Hypothesenraum* wird dieses in den Zielzustand, die überprüfbare Hypothese, überführt. Das *Testen der Hypothesen* beginnt nun mit der Auswahl bzw. dem Design und der Durchführung geeigneter Experimente im Experiment-Raum. Output des Experiment-Raumes sind die experimentellen Ergebnisse, welche im letzten Schritt – der *Analyse der Evidenzen* – verifiziert oder falsifiziert werden (Hammann et al. 2007). Die Hauptkomponenten des SDDS-Modells sind der folgenden Abbildung 18 zu entnehmen.

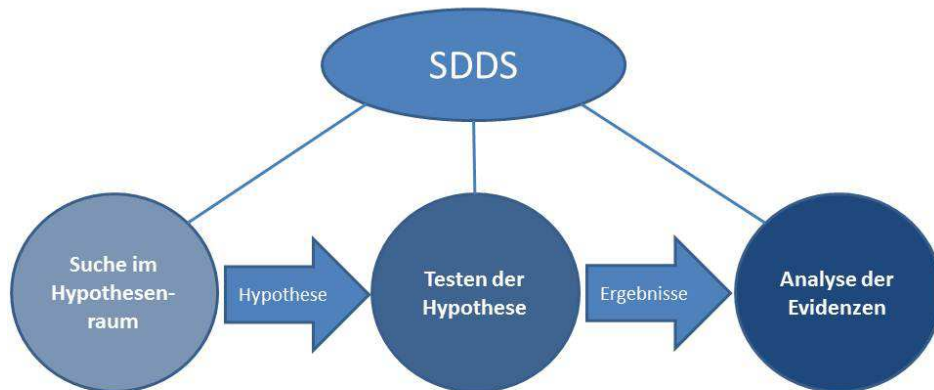


Abbildung 18: Die drei Hauptkomponenten des SDDS-Modells (verändert nach Klahr 1988 & Hammann et al. 2007)

Klahr und Dunbar (1988) führten verschiedene Versuche durch, um die im Modell postulierten Prozesse zu bestätigen. Im Rahmen einer Studie wiesen sie Erwachsene in die grundlegende Bedienung des programmier- und fahrbaren Roboters BigTrak ein. Die Bedeutung der Bedienungstaste *repeat* wurde den Versuchsteilnehmern allerdings vorenthalten. Mit dem Hinweis, dass die Funktion dieser Taste im Programm nur einmal vorkommen kann und dass ihr ein Zahlenwert zuzuordnen ist, war es Aufgabe der Versuchsteilnehmer, die Funktion unter verbaler Äußerung ihrer Gedankengänge, selbst herauszufinden. Im Rahmen dieser Studie konnten Klahr und Dunbar (1988) zwei Experimentiertypen identifizieren. Während die Theoretiker mehr Zeit im Hypothesenraum verbrachten, indem sie ihr Vorwissen für die gezielte Entwicklung von geeigneten Hypothesen einsetzten, entwickelten die Experimentierer geeignete Hypothesen erst beim Experimentieren selbst.

1.2.4.5.2 Kompetenzentwicklungsmodell nach Hammann (2004)

Auf Grundlage des SDDS-Modells von Klahr & Dunbar (1988) entwarf Hammann (2004) ein Modell zur zeitlichen Entwicklung der Experimentierkompetenz im Verlauf der Sekundarstufe I. Hier differenziert er die drei Teilkomponenten des SDDS-Modells unter Einbeziehung der

Kompetenzniveaus naturwissenschaftlicher Grundbildung nach Bybee (2002) und beschreibt, welches Kompetenzniveau in welcher Klasse zu erreichen ist. Exemplarisch ist diese Differenzierung in Tab. 4 für die Suche im Hypothesenraum ausgeführt.

Tabelle 4: Kompetenzstufen bei der Suche im Hypothesenraum (verkürzt nach Hammann 2004)

Stufe	Kompetenzniveau	Nähere Beschreibung	Alter
1	Keine Hypothesen beim Experimentieren	Es wird ohne Hypothesen experimentiert, das heißt, Lernende führen Experimente aus, ohne Vermutungen über Ursache Wirkungs-Beziehungen zu haben [...]	Primarstufe
2	Unsystematische Suche nach Hypothesen	Es wird hypothesengeleitet experimentiert, jedoch werden nicht alle Hypothesen herangezogen, die für die Beantwortung einer Fragestellung beachtet werden müssten [...]	Klasse 5
3	Systematische Suche nach Hypothesen	Es werden multiple Hypothesen gebildet, die logisch aufeinander bezogen werden. Probleme bereitet jedoch die Hypothesenrevision.	Klasse 5
4	Systematische Suche nach Hypothesen und erfolgreiche Hypothesenrevision	Wie Stufe 3, jedoch gelingt auch die Hypothesenrevision in Situationen, in denen alle bereits getesteten Hypothesen falsifiziert wurden.	Klasse 7

Da es sich hier nicht um ein neues Struktur-, sondern um ein auf dem Modell von Klahr & Dunbar (1988) aufbauendes Kompetenzentwicklungsmodell handelt, wird für detailliertere Ausführungen auf Hammann (2004) verwiesen.

1.2.4.5.3 Strukturmodell zum wissenschaftlichen Denken (Mayer 2007)

Auch im Strukturmodell zum wissenschaftlichen Denken nach Mayer (2007, Abb. 19) nimmt die Experimentierkompetenz eine zentrale Rolle ein. In diesem Modell wird wissenschaftliches Denken als domänenspezifischer, wissensbasierter und komplexer Problemlöseprozess beschrieben. Dieser ist durch spezifische Prozeduren charakterisiert und greift auf inhaltliches und methodisches Wissen zurück (Mayer et al. 2008). In der Betonung des Problemlösecharakters des naturwissenschaftlichen Denkens gleicht Mayers (2007) Modell dem SDDS-Modell nach Klahr (2000). Wie Abbildung 19 darstellt, berücksichtigt Mayer (2007) hier nicht nur die prozessbezogenen Variablen des naturwissenschaftlichen Erkenntniswegs, sondern auch personenbezogene Einflüsse auf die Kompetenz des naturwissenschaftlichen Denkens (*scientific reasoning*).

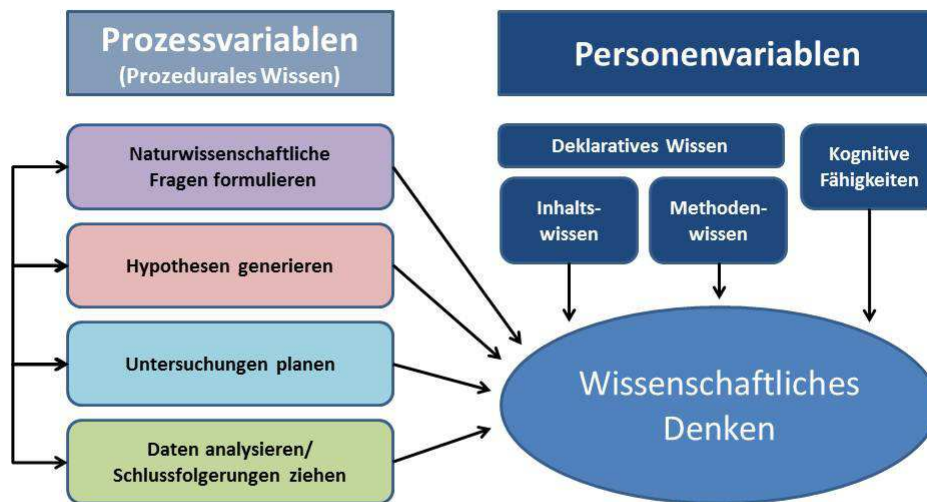


Abbildung 19: Kompetenzmodell naturwissenschaftlichen Denkens (nach Mayer 2007, Mayer et al. 2008; die Kompetenzniveaus der vier farblich kodierten Teilkompetenzen finden sich in Tab. 5)

Unter den Prozessvariablen finden sich die Elemente des naturwissenschaftlichen Erkenntnisweges gegliedert in die Teilkompetenzen Naturwissenschaftliche Fragen formulieren, Hypothesen generieren, Untersuchungen planen und Daten analysieren/Schlussfolgerungen ziehen. Diese Teilkompetenzen werden als eigenständige Konstrukte betrachtet, die in gegenseitiger Wechselbeziehung stehen. Im Unterschied zum SDDS-Modell definiert Mayer (2007) hier vier und nicht drei Hauptkomponenten des Erkenntnisweges. In ihrer Studie bestätigen Mayer et al. (2008) die im Modell Mayers (2007) postulierte Vierdimensionalität der prozeduralen Variablen. Für jede der vier Teilkompetenzen des naturwissenschaftlichen Erkenntnisweges werden die fünf in Tabelle 5 beschriebenen Kompetenzniveaus formuliert (Mayer et al. 2008). Wie Möller et al. (2007) feststellen, nimmt das Kompetenzniveau jeder Teilkompetenz mit steigender Jahrgangsstufe zu.

Während sich im SDDS-Modell die Durchführung von Experimenten der Komponente des Experimentiersuchraumes zuordnen ließe, ist dieser praktische Teilaspekt des Experimentierens nicht expliziter Gegenstand des Modells zum naturwissenschaftlichen Denken. Bedeutender Einfluss wird allerdings den Personenvariablen zugesprochen. Naturwissenschaftliches Denken erfordert demnach, auf das individuell vorhandene deklarative Wissen zurückgreifen zu können. Dieses setzt sich zusammen aus dem Inhaltswissen, also dem naturwissenschaftlichen Faktenwissen sowie dem Methodenwissen. Darunter wird das Wissen über die Methoden naturwissenschaftlicher Erkenntnisgewinnung, also beispielsweise die Kenntnis von Experimentier- und Datenauswertungsverfahren verstanden (Mayer et al. 2008). Auch Hammann (2004) stellt fest, dass sich neben dem mangelnden methodischen Wissen über das Experimentieren auch Defizite im inhaltlichen Vorwissen begrenzend auf experimentelle Fähigkeiten auswirken. Wie bereits im

Rahmen der Studien zum SDDS-Modell erwähnt wurde, ist die kognitive Disposition also nicht unerheblich für den naturwissenschaftlichen Erkenntnisprozess.

Tabelle 5: Kompetenzniveaus des wissenschaftlichen Denkens (verändert nach Mayer et al. 2008; die farbliche Kodierung der vier Teilkompetenzen wurde aus Abb. 19 übernommen)

Niveau	Beschreibung	Teilkompetenzen
V	Selbständiges Lösen von offenen Problemen	Eigene naturwiss. Fragen zur Problemlösung formulieren
		Alternative Hypothesen berücksichtigen
		Untersuchungsmethoden, Genauigkeit, Fehler abwägen
		Alternative Deutungen und Erklärungen in Betracht ziehen
IV	Elaborierte Untersuchung allgemeiner Zusammenhänge	Naturwiss. Fragen nach einem verallgemeinernden Zusammenhang formulieren
		Generalisierte Hypothesen formulieren
		Stichprobe, Messwiederholung und Versuchsdauer berücksichtigen
		Deutungen hinsichtlich Sicherheit der Daten und Reichweite reflektieren
III	Kontrollierte Untersuchung auf der Basis von Konzeptverständnis	Naturwiss. Fragen zu einem quantitativen Zusammenhang von Variablen formulieren
		Hypothesen auf der Basis von Konzeptverständnis begründen
		Kontrollvariablen berücksichtigen
		Erklären von Daten auf der Basis von Konzeptverständnis
II	Untersuchung von Zusammenhängen	Naturwiss. Fragen nach Zusammenhang zweier Variablen formulieren
		Hypothesen mit Alltagsanalogien begründen
		Veränderte und zu messende Variable in Beziehung setzen
		Schlussfolgerungen aus Daten ziehen
I	Untersuchung eines Faktors	Einfache naturwiss. Fragen auf Phänomenebene stellen
		Einfache, testbare Hypothese generieren
		Eine Variable identifizieren
		Beobachtung/Daten wiedergeben

1.2.4.5.4 Modell experimenteller Kompetenz (Schreiber et al. 2009)

Sehr konkret formulieren Schreiber et al. (2009) die Teilkompetenzen der Experimentierfähigkeit. Ähnlich wie im SDDS-Modell finden sich hier drei zentrale Komponenten (Abb. 20). Während im SDDS-Modell die Hypothesenbildung und die Interpretation der Ergebnisse im Vordergrund stehen, sehen Schreiber et al. (2009) allerdings die Durchführung von Experimenten im Fokus ihres Modells. Die vorausgehende Planung und die Auswertung der experimentellen Ergebnisse werden in Orientierung an Klahr (2000) und Hammann (2004) mit erfasst. Die drei in diesem Modell ausgeführten Teilkompetenzen geben einen umfassenden Überblick über das Konstrukt der Experimentierkompetenz und sollen daher im Anschluss erläutert werden. Schreiber et al. (2009) betonen ausdrücklich, dass die Teilkompetenzen im Prozess des Experimentierens nicht zwingend in linearer Abfolge gefordert sind. Sie gehen vielmehr davon aus, dass die Merkmale in einem

Experimentierprozess in verschiedenen Reihenfolgen, iterativ und nicht immer vollständig auftreten können (Schreiber et al. 2009). Nachfolgend wird ein Überblick über die Bedeutung der drei Teilkompetenzen gegeben.



Abbildung 20: Modell experimenteller Kompetenz (nach Schreiber et al 2009)

1.2.4.5.4.1 Planungskompetenz

Am Anfang eines Forschungsprozesses steht eine Problemstellung, häufig in Form eines unbekannten Phänomens. In vielen Fällen ist der Forscher zunächst gefordert, eine Fragestellung zur Klärung des Phänomens zu entwickeln. Wie Roth (1995) feststellt, erhöht sich die Fähigkeit von Schülern, gezielte wissenschaftliche Fragen zu formulieren, mit zunehmender Vertrautheit mit einem spezifischen Kontext. Die Formulierung einer eigenen Fragestellung erfordert daher bereits eine, in Abhängigkeit vom Vorwissen mehr oder weniger umfangreiche Auseinandersetzung mit dem Problemgegenstand. Viele Forschungsprojekte basieren auch auf bereits klar definierten, vorgegebenen Fragestellungen. Sowohl vorgegebene als auch selbst entwickelte Fragestellungen eröffnen die Möglichkeit, Erwartungen und schließlich gut begründete Hypothesen über das Phänomen zu formulieren (Schreiber et al. 2009). Kommt im SDDS-Modell (Klahr & Dunbar 1988) der Entwicklung eigener Fragestellungen eine hohe Bedeutung zu, so betonen Schreiber et al. (2009) die – gerade im Rahmen von Lehr-Lernsituationen – hohe Relevanz der Beantwortung vorgegebener Fragestellungen. Diese Fokussierung auf vorgegebene Fragen erscheint insbesondere für außerschulische Unterrichtsangebote bedeutsam, da hier häufig der zeitliche Rahmen die selbstständige Entwicklung einer Fragestellung limitiert. Der Fähigkeit, anhand der Hypothesen zunächst gedanklich und dann real einen Versuchsplan zu entwerfen, ordnen Schreiber et al. (2009) in ihrem Modell eine Zwischenstellung zwischen der Planung und Durchführung von Experimenten zu.

1.2.4.5.4.2 Durchführungskompetenz

Bei der praktischen Umsetzung des Versuchsplanes liegen die Schwierigkeiten oft im Detail. Bereits die Zusammenstellung der für den Versuch geeigneten Geräte und Chemikalien macht eine Kenntnis der lokal vorhandenen und der Vielfalt der im Laborbedarf erhältlichen Geräte erforderlich. Während der Aufbau einer Versuchsanordnung bei bereits gut etablierten Versuchen oder der Verfügbarkeit kompletter Experimentier-Kits gegebenenfalls leicht fällt, steht der Neukonstrukteur einer Versuchseinrichtung häufig vor dem Problem der mangelnden Kompatibilität der Einzelkomponenten. Letztlich wird vom Experimentator die Kompetenz gefordert, Messungen durchzuführen und diese in angemessener Form zu dokumentieren (Schreiber et al. 2009). Die in dieser Teilkompetenz geforderten Fähigkeiten spielen eine besondere Rolle für außerschulische Experimentierlabore, da hier der Schwerpunkt der Tätigkeiten in der Regel in der praktischen Durchführung von Experimenten liegt.

1.2.4.5.4.3 Auswertungskompetenz

Häufig funktioniert ein neuer Versuch nicht immer wie geplant. Unerwartete Probleme machen die Fehlersuche und die Vornahme von Modifikationen zur Optimierung des Versuchsablaufes erforderlich. Die Kompetenz im Umgang mit solchen Problemen und Fehlern findet sich im Modell von Schreiber et al. (2009) als Schnittstelle zwischen der Durchführung und Auswertung von Versuchen wieder. Die im finalen Versuch dokumentierten Messdaten müssen letztlich in angemessener Weise aufbereitet und verarbeitet werden. Insbesondere bei größeren Datenmengen erfordert dies vom Experimentator die Beherrschung von Verfahren der deskriptiven (Übersichtliche Darstellung der Daten in Tabellen mit Hilfe von Kennwerten und Grafiken) und schließenden Statistik (Durchführung geeigneter statistischer Tests). Erst die abschließende Interpretation der Daten unter Einbezug von eigenem Vorwissen und Literatur ermöglichen die Verifizierung oder Falsifizierung der anfänglich aufgestellten Hypothesen. Die zunehmende Vertrautheit der Schüler mit dem spezifischen Kontext verändert ihre Fähigkeit, die Ergebnisse eines Experimentes zu deuten. Stehen zunächst noch simple, unzulängliche Erklärungen im Vordergrund, so sind Schüler bei längerer Beschäftigung mit dem Kontext zunehmend befähigt, auch komplexe Beziehungen in ihren Daten aufzudecken (Roth 1995). Wie schon die Planung von Experimenten, so erfordert auch eine umfassende Datenauswertung einen umfangreicheren Zeiteinsatz, als er häufig bei außerschulischen Lernangeboten gewährleistet werden kann. Dementsprechend müssen ausstehende Auswertungen im dem auf den Laborbesuch folgenden Unterricht ergänzt oder weitergeführt werden.

Aufgrund der sehr konkret ausformulierten Teilkompetenzen bietet sich das Modell experimenteller Kompetenz nach Schreiber et al. (2009) als Grundlage für die Entwicklung von Testinstrumenten zur

Erfassung experimenteller Fähigkeiten an. Damit stellt es ein Modell hoher Praxisrelevanz dar. Einflüsse dieses Modells sind auch im Spinnenmodell experimenteller Kompetenz (Nawrath et al. 2011), welches nachfolgend vorgestellt wird, zu finden.

1.2.4.5.5 Spinnenmodell experimenteller Kompetenz (Nawrath et al. 2011)

Im Rahmen des vom IPN Kiel koordinierten Forschungsprojektes *komdif* (Kompetenzentwicklungsmodelle als Basis für eine diagnosegestützte Förderung von Schülerinnen und Schülern in der Primarstufe und Sekundarstufe I) wurde am Institut für Didaktik der Naturwissenschaften an der Universität Bremen (IDN) in Kooperation mit Lehrkräften des Schulversuchs „alles>>können“ das Spinnenmodell entwickelt (Abb. 21). Dieses auf den Modellen von Hammann (2004), Mayer (2007) und Schreiber et al. (2009) basierende Modell soll in der Schulpraxis Lehrkräften dabei helfen, experimentellen Unterricht kompetenzorientiert zu gestalten und die experimentellen Fähigkeiten der Schüler zu diagnostizieren (Maiseyenka et al. 2011).

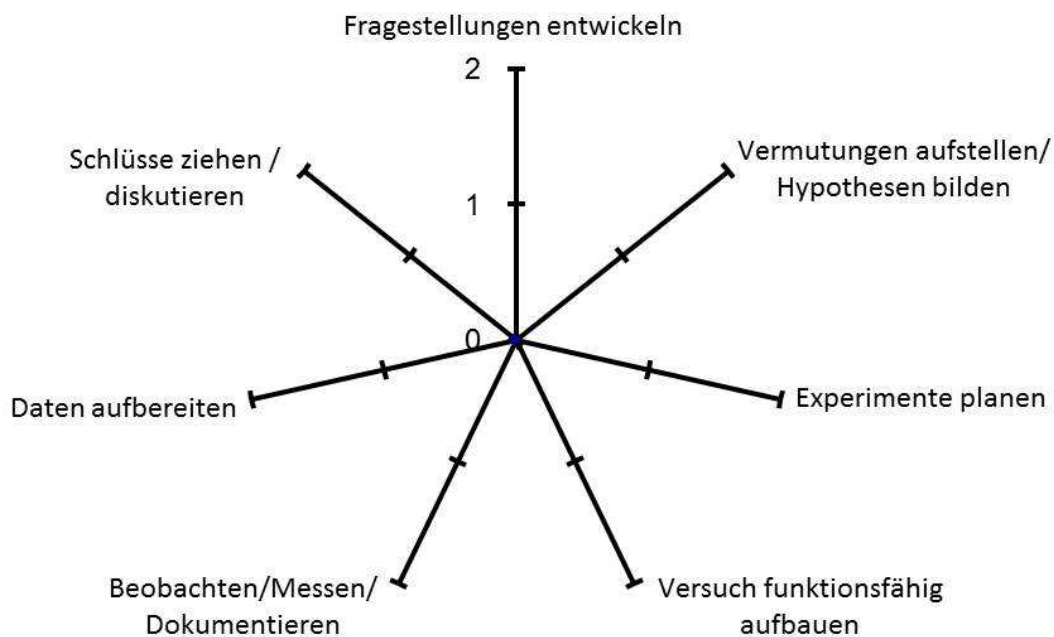


Abbildung 21: Teildimensionen experimenteller Kompetenz im Spinnenmodell (Graduierung der Teildimensionen bei der Planung von Unterricht 0: unwichtig, 1: bedeutsam, 2: besonders wichtig; Kompetenzstufen bei der Diagnose: 0: niedriges Niveau, 1: mittleres Niveau, 2: hohes Niveau; Nawrath et al. 2011)

Die Graduierung der sieben Teildimensionen von null bis zwei dient bei der Planung und Reflexion naturwissenschaftlichen Unterrichts als Maß für die eingeschätzte Relevanz jeder Teildimension im geplanten Unterrichtsexperiment (Abb. 21). Als Diagnoseinstrument entspricht die Graduierung der von den Schülern beim unterrichtlichen Experimentieren erreichten Kompetenzausprägungen. Erste Ausformulierungen der Niveaustufen liegen aktuell bereits vor (Maiseyenka et al. 2011).

Wie zuvor beschrieben geht die Bedeutung des Spinnenmodells über ein reines Strukturmodell hinaus, da es gleichzeitig als Instrument zur Messung von Experimentierfähigkeit Verwendung finden soll. Wie der folgende Abschnitt zeigt, ist das Interesse an der Messung von Experimentierkompetenz nicht neu.

1.2.4.6 Ansätze zur Erfassung experimenteller Kompetenz

Entsprechend der Bildungsstandards sollen Schüler im Kompetenzbereich Erkenntnisgewinnung unter anderem Experimentierkompetenz entwickeln (KMK 2005). Daher ist es notwendig, für die Diagnose experimenteller Kompetenz geeignete Messinstrumente zu entwickeln (Schreiber et al. 2009). Kompetenzmodelle bilden die Grundlage zur systematischen Erfassung von Kompetenzen. Erst die konkrete Beschreibung der Experimentierkompetenz macht dieses Konstrukt operationalisierbar und damit vollständig messbar. Bisherige Tests zur Erfassung der Experimentierkompetenz greifen fast immer auf Experimentaltests unter Durchführung von Realexperimenten (*practical test* oder *laboratory performance test*) oder schriftliche Wissenstests (*paper-and-pencil test*) zurück (Hammann et al. 2007, Schreiber et al. 2009). Vertreter traditioneller schriftlicher Testverfahren sind der *Test of Science Processes* (Tannenbaum 1971) oder der *Test of Enquiry Skills* (TOES, Fraser 1980). Der *Test of Science Processes* besteht aus 96 Multiple-choice-Items und bemüht sich, die Fähigkeiten von Siebt-, Acht-, und Neuntklässlern in den acht Subskalen *Observing*, *Comparing*, *Classifying*, *Quantifying*, *Measuring*, *Experimenting*, *Inferring* und *Predicting* zu messen. Die KR-20²-Reliabilität der aus zehn Items bestehenden Subskala *Experimenting* betrug in Abhängigkeit von der Klassenstufe und dem untersuchten Einzugsgebiet zwischen 0,43 und 0,45. Für die Gesamtheit der Skalen ist eine Bearbeitungsdauer von maximal 73 Minuten angegeben (Tannenbaum 1971). Im TOES werden neun Subskalen mit insgesamt 87 Items eingesetzt, von denen insbesondere die Subskala *Design of experimental procedures* (KR-20 zwischen 0,50 bis 0,61 in Abhängigkeit von der Klassenstufe) und die Subskala *Conclusion and generalizations* (KR-20 zwischen 0,62 bis 0,69 in Abhängigkeit von der Klassenstufe) die Erfassung von Experimentierfähigkeit anstreben. Die Bearbeitungsdauer für jede der neun Subskalen beträgt etwa zehn Minuten (Fraser 1980). Ein weiteres schriftliches, auf Grundlage des SDDS-Modells (vgl. Punkt 1.2.4.5.1.) entwickeltes Testverfahren setzen Hammann et al. (2007) ein. Dieser Kompetenztest umfasst 24 Items (in jeder Dimension des SDDS-Modells acht Items) und wird kombiniert mit einem Wissenstest aus 31 Items eingesetzt. Der Test richtet sich an die Klassenstufe fünf bis sechs und ist innerhalb von 45 Minuten durchführbar. Im Vergleich mit den älteren Tests fallen die Reliabilitäten der drei Subskalen mit einem Cronbachs α -Wert zwischen 0,69

² Die Kuder-Richardson-Formel 20-Reliabilität (KR-20) ist ein vor der Einführung von Cronbachs α häufig angegebenes Maß für die Reliabilität bei Vorliegen dichotomer Variablen. Bei dichotomen Items wird in SPSS automatisch die KR-20-Reliabilität berechnet (vgl. Bühl 2010).

und 0,78 deutlich besser aus. Allerdings weist die faktorenanalytische Betrachtung darauf hin, dass das Messinstrument lediglich zwei, statt drei Dimensionen abbildet. Grube (2011) entwickelte einen schriftlichen Leistungstest zur Erfassung der vier Teilkompetenzen Fragestellung entwickeln, Hypothesen ableiten, Experimente planen und Experimente auswerten. Jede Subskala wird in sechs Items erfasst. Die Skalenreliabilitäten liegen zwischen Cronbachs $\alpha=0,49$ und 0,75. Im Unterschied zu den bisher vorgestellten Tests wählt Grube (2011) ein offenes Antwortformat. Für jede Subskala werden fünf Kompetenzniveaus definiert, auf die die Antworten der Schüler hin zu analysieren sind.

Ein großer Vorteil der rein schriftlichen Erhebungsverfahren ist in ihrer Ökonomie zu sehen, da praktische Experimentalaufgaben in der Regel einen hohen Durchführungsaufwand erfordern. Allerdings wird häufig angezweifelt, dass derartige *paper-pencil*-Tests valide genug sind, um das ganze Spektrum experimenteller Fähigkeiten zu erfassen. Insbesondere die Kompetenz, Experimente durchzuführen ist so kaum überprüfbar. Praktische Testverfahren wählen daher einen anderen Weg. Zu differenzieren ist hier zwischen prozessbezogenen Experimentaltests, bei denen Handlungsanalysen bei der Tätigkeit des Experimentierens vorgenommen werden und produktbezogenen Testverfahren, bei denen die Auswertung aufgrund schriftlich fixierter Arbeitsergebnisse der Schüler erfolgt (Schreiber et al. 2009). Ein produktbezogenes Testverfahren setzte Tamir (1974) ein, um die Experimentierfähigkeit von Zwölftklässlern bei einer praktischen Abschlussprüfung an israelischen *high schools* zu erfassen. Jeder Schüler erhält im Rahmen dieser Prüfung ein *examination paper*, welches Anweisungen zur Durchführung eines etwa neunzigminütigen Versuchs enthält. Sämtliche Versuchsmaterialien werden zur Verfügung gestellt, so dass anhand des *examination paper* weitgehend selbstständig gearbeitet werden soll. Die Aufgaben im *examination paper* sind im Laufe des Versuches zu bearbeiten. Im Anschluss werden die Resultate der Aufgaben unter den Gesichtspunkten praktische Handhabung, Selbstständigkeit, Beobachtungsgabe, experimentelles Design, Auswertung und Kommunikation sowie logisches Denken zur Beurteilung der praktischen Fähigkeit ausgewertet. Die Reliabilität der erfassten Skalen schwankt zwischen Cronbachs $\alpha=0,56$ und 0,69. Ben-Zvi et al. (1977) setzen neben einem Wissenstest eine prozessbezogene Dokumentation ein, um die Experimentierfähigkeit von 15jährigen Schülern zu erfassen. Dazu wird den Teilnehmern eine Problemstellung geliefert, welche es durch Planung und Durchführung eines Experiments zu lösen gilt. Das praktische Vorgehen wird von einem Gutachter verfolgt und anhand einer Checkliste mit einem Gesamtpunktwert beurteilt. In Abhängigkeit des beobachteten Versuchs liegt die Reliabilität dieses Tests zwischen Cronbachs $\alpha=0,53$ und 0,70. Andere prozessbezogene Erhebungsverfahren setzen Videoanalysen mit anschließender, kategoriebasierter Auswertung zur Beurteilung von Experimentiertätigkeiten ein (vgl. Neumann 2004, Tesch & Duit 2004). In Anbetracht des hohen Material- und Analyseaufwands, den die prozessbezogene Auswertung von Realexperimenten erfordert, entwickelten Schreiber et al.

(2009) ein auf computersimulierten Experimenten basierendes Erhebungsverfahren. In Simulationsbaukästen erhalten die Schüler am PC die Möglichkeit, Experimente selbst aufzubauen und durchzuführen. Die Datenerhebung erfolgt über Protokolle, die von den Probanden zu führen sind, Bildschirmaufzeichnungen und Navigationsdaten, welche die Rekonstruktion der praktischen Tätigkeiten erlauben sollen. Studien zum Vergleich dieses Erhebungsverfahrens mit den Beobachtungsdaten aus Realexperimenten und schriftlichen Wissenstests stehen bevor, so Schreiber et al. (2009). Shavelson et al. (1999) stellen in einem Vergleich verschiedener Testverfahren zur Erfassung der praktischen Experimentierfähigkeit eine deutliche Abhängigkeit des Testergebnisses vom Messverfahren fest. Besonders hohe Korrelationen finden sie zwischen den Ergebnissen der Live-Beobachtung von Experimenten (*direct observations*) und der Auswertung von stichwortartigen Kurzprotokollen (*notebook*), welche die Schüler während des Experimentierens selbst fertigten. Dem gegenüber bestehen nur relativ geringe Korrelationen zwischen den *direct observations* und schriftlichen *paper-and-pencil*-Tests.

1.2.5 Fachdidaktische Fragestellungen

Wie im vorangegangenen Abschnitt gezeigt wurde, existieren verschiedene Ansätze zur Erfassung der Experimentierfähigkeit. Häufig wird schriftlichen Wissenstests eine mangelnde Validität zur Erfassung der Experimentierkompetenz nachgesagt. Insbesondere die Kompetenz, Experimente durchzuführen wird in diesem Testformat in der Regel kaum berücksichtigt. Praktische Experimentaltests erscheinen dagegen gerade für den großflächigen Einsatz im Rahmen von Lernstandserhebungen zu aufwändig (Schreiber et al. 2009). Einen interessanten Ansatz verfolgt Brandt (2005), welcher in seiner aus fünf Items bestehenden Fragebogenskala „*teutolab* chemie-Selbstkonzept der Begabung“ nicht unmittelbar die Experimentierkompetenz misst, sondern die subjektiv eingeschätzte Experimentierfähigkeit erhebt. Dies geschieht allerdings ohne Berücksichtigung der Teilkompetenzen des Experimentierens. Seine Items wie „Solches Experimentieren wie im *teutolab* liegt mir nicht besonders“ stützen sich nicht auf ein Modell experimenteller Kompetenz, sondern wurden durch Umformulierung von Items zur Erfassung des fachlichen Selbstkonzeptes der Begabung konstruiert. Dennoch scheint die Grundidee vielversprechend, denn wie im Abschnitt 1.2.3.3. vorgestellt wurde, stellt die selbst eingeschätzte Kompetenz einen möglichen Prädiktor für tatsächliches Verhalten dar. Diese Grundidee findet in der vorliegenden Forschungsarbeit Berücksichtigung in der Entwicklung eines modellbasierten Messinstrumentes zur Erfassung der experimentbezogenen Selbstwirksamkeit. Bezogen auf die Entwicklung dieses neuen Messinstrumentes ist folgende Fragestellung von Relevanz:

Lassen sich die drei Teildimensionen des Modells experimenteller Kompetenz nach Schreiber et al. (2009) in Likert-skalierten Items operationalisieren und in ein reliables Messinstrument zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung überführen?

Im Abschnitt 1.2.3.3. wurde die Selbstwirksamkeit als potentieller Prädiktor für Verhalten vorgestellt. Von Interesse ist daher für die vorliegende Studie auch die folgende Überlegung:

Liefert die experimentbezogene Selbstwirksamkeitserwartung einer Person Hinweise auf ihre tatsächliche Kompetenz beim Experimentieren?

Umfassende Studien zur experimentbezogenen Selbstwirksamkeitserwartung von Schülern liegen bisher kaum vor. Die vorliegende Studie untersucht den Einfluss der außerschulischen Lernumgebung BeLL Bio auf dieses Konstrukt.

Fraglich ist, wie Schüler ihre Experimentierkompetenz in praktischen Experimentierkursen im Schülerlabor einstufen und wie sich der Laborbesuch auf die experimentbezogene Selbstwirksamkeitserwartung auswirkt.

Das fachliche Selbstkonzept und das Interesse einer Person gelten als bedeutsame Merkmale der persönlichen Disposition sowie als wichtige Größen mit Einfluss auf Lernerfolg und Leistungsmotivation. In einigen Schülerlaboren konnten bereits Einflüsse einmaliger Laborbesuche auf diese Konstrukte nachgewiesen werden (Brandt 2005, Pawek 2009). Auch das neu etablierte BeLL Bio formuliert in seinen Zielen, Interesse für die Naturwissenschaft Biologie sowie Selbstvertrauen im Umgang mit biologischen Sachverhalten erzeugen zu wollen.

Welchen Einfluss hat ein einmaliger Laborbesuch im Lernort BeLL Bio auf das aktuelle Interesse und das Selbstkonzept?

Als außerschulischer Lernort beabsichtigt das Schülerlabor BeLL Bio nicht nur die affektive Aktivierung von Schülern. Die Teilnahme an den angebotenen Laborkursen soll auch ein spannendes, nachhaltiges Lernerlebnis bieten. Eine Überprüfung des durch einmalige Schülerlaborbesuche bewirkten Lernerfolges wird nur in sehr wenigen Studien vorgenommen, so dass diesem Aspekt in der vorliegenden Studie besondere Relevanz zukommt.

Es gilt daher zu prüfen, ob die im BeLL Bio angebotenen Laborkurse zu einem nachhaltigen kognitiven Wissenserwerb beitragen.

Um ein ganzheitlicheres Bild der Einflussgrößen am Lernort BeLL Bio zu erreichen, erscheint eine singuläre Betrachtung der eben genannten Fragestellungen nicht ausreichend. Eine umfassendere Beschreibung verlangt die gleichzeitige Berücksichtigung möglichst vieler Einflussgrößen in Form

eines Hypothesennetzes. In diesem werden einzelne, theoretisch begründete Hypothesen zu einem komplexen Ursache-Wirkungsgefüge zusammengefasst (Weiber & Mühlhaus 2010). Ein solches Hypothesennetz affektiver Einflussgrößen am Lernort BeLL Bio wird hier theoriegeleitet aus den in den Abschnitten 1.2.2 und 1.2.3 vorgestellten Einflussgrößen hergeleitet, um im Rahmen der Auswertung seine Gültigkeit prüfen zu können. Es orientiert sich am Modell der Interessenförderung durch Schülerlabore nach Pawek (2009) und verändert und erweitert dieses um die Variablen der experimentbezogenen Selbstwirksamkeitserwartung, des Berufsinteresses und um eine differenziertere Betrachtung des Einflusses von Fach- und Sachinteresse (Abb. 22).

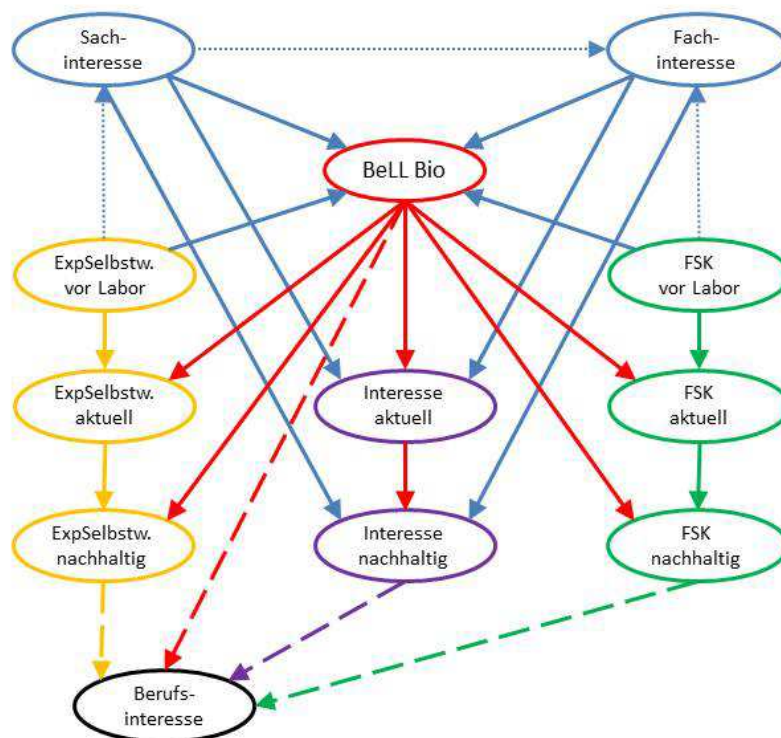


Abbildung 22: Hypothesennetz zentraler affektiver Einflussgrößen im BeLL Bio (ExpSelbstw.=Experimentbezogene Selbstwirksamkeitserwartung; FSK=Fähigkeitsselbstkonzept)

Im Zentrum des in Abbildung 22 dargestellten Hypothesennetzes stehen die den Laborbesuch repräsentierenden Einflussgrößen, hier als rote Ellipse symbolisiert. Dazu zählen die vorgestellten Laborvariablen, also der Alltagsbezug, die Authentizität, die Atmosphäre, die Betreuerwahrnehmung, die Zusammenarbeit, die Transparenz der Versuchsanleitungen sowie der empfundene Druck. Wie Schüler den Laborbesuch wahrnehmen, hängt von ihrer persönlichen Disposition ab. Pawek (2009) fasst das vor dem Laborbesuch bestehende naturwissenschaftliche Interesse sowie das Interesse am Experimentieren als Sachinteresse zusammen. Er stellt einen deutlichen Einfluss dieses Sachinteresses auf die Wahrnehmung des Laborbesuchs fest. Daher wird im oben abgebildeten Hypothesennetz ebenfalls ein Einfluss des dispositionalen Fachinteresses sowie des Sachinteresses auf die Laborwahrnehmung angenommen. Dieser wird symbolisiert durch die von beiden

Interessenkonstrukten ausgehenden, zum Konstrukt BeLL Bio weisenden, blauen Pfeile. Die Trennung von Fach- und Sachinteresse erscheint sinnvoll, da ein geringes Fachinteresse nicht notwendigerweise mit einem geringen Sachinteresse einhergehen muss (Hoffmann et al. 1998, Wodzinski 2002). Da das Experimentieren elementarer Bestandteil des Biologieunterrichts ist, wird im beschriebenen Hypothesensystem angenommen, dass sich das Sachinteresse am Experimentieren auf das Fachinteresse auswirkt. Ausgehend von der Interessentheorie (vgl. Abschnitt 1.2.3.4) wird die Hypothese aufgestellt, dass sowohl das persönliche dispositionale Interesse als auch die Interessantheit der Lernumgebung BeLL Bio Einfluss auf das aktuelle Interesse beim Laborbesuch hat. Dies sollte nicht nur kurzfristig, sondern auch nachhaltig zu beobachten sein, wenn der Laborbesuch bzw. der ihm folgende Schulunterricht ausreichend *hold*-Komponenten (Mitchell 1993) bietet, um für die Schüler persönlich bedeutsam zu werden (vgl. auch Guderian et al. 2006b). Als weitere, bereits vor dem Laborbesuch bestehende dispositionale Größen, wird ein Einfluss des biologischen Fähigkeitsselbstkonzeptes sowie der experimentbezogenen Selbstwirksamkeitserwartung auf die Wahrnehmung des BeLL Bio-Besuches angenommen. Schüler mit einer niedrigeren Selbstwirksamkeitserwartung werden in der Experimentiersituation BeLL Bio voraussichtlich zögerlicher agieren als diejenigen mit einer hohen Selbstwirksamkeitserwartung (Asendorpf 2007). Sie werden beim Experimentieren dementsprechend weniger aktiv mit ihren Mitschülern zusammenarbeiten, aber auch zögerlicher mit den Betreuern in Kontakt treten. Somit ist von einem positiven Einfluss einer hohen dispositionalen experimentbezogenen Selbstwirksamkeitserwartung auf die Wahrnehmung des Laborbesuches auszugehen. Da die Selbstwirksamkeit Element des Selbstkonzeptes ist (Woolfolk 2008), ist auch für dieses im Hypothesensystem ein ähnlicher Zusammenhang zu formulieren. Dementsprechend sollten beide Konstrukte positiv miteinander korrelieren. Bei Pawek (2009) bestätigt sich der Einfluss des dispositionalen Selbstkonzeptes auf die Laborwahrnehmung allerdings nicht.

Wie eben beschrieben wird eine Person, die sich in Bezug auf das Experimentieren als selbstwirksam einschätzt, wahrscheinlich weniger zögerlich in einer Experimentiersituation agieren. Sie wird sich also im Sinne der Person-Gegenstands-Theorie des Interesses intensiver mit dem Interessengegenstand auseinandersetzen. Daher wird ein Einfluss der Selbstwirksamkeit auf das Sachinteresse angenommen. Aus analogen Gründen wird im Hypothesennetz ein Einfluss des biologischen Fähigkeitsselbstkonzeptes auf das Fachinteresse postuliert. Innerhalb verschiedener Studien wurde festgestellt, dass das Selbstkonzept am Lernort Labor (Engeln 2004) bzw. am Lernort Praktikum (Neumeyer & Mikelskis 2009) deutlich schwächer mit dem aktuellen Interesse korreliert als das Fach- und Sachinteresse. Auch Pawek (2009) stellt in seinem Modell keinen direkten Einfluss des dispositionalen Selbstkonzeptes auf das aktuelle Interesse nach dem Laborbesuch fest. Aus

diesem Grund wurde im vorliegenden Hypothesenkonstrukt kein Einfluss des Fähigkeitsselbstkonzepts und der Selbstwirksamkeit auf das aktuelle Interesse angenommen.

Selbstwirksamkeit und somit auch das eng dazu in Beziehung stehende Fähigkeitsselbstkonzept sind keine statischen Konstrukte (Woolfolk 2008). Das didaktische Konzept der BeLL Bio-Kurse berücksichtigt die Lernvoraussetzungen der Schüler und möchte ein ideales Herausforderungsniveau bieten, welches fachliche und sachliche Kompetenzerlebnisse ermöglicht. Insbesondere die Person des Betreuers sollte durch konstruktive Rückmeldungen bedeutsamen Einfluss auf dieses Kompetenzerlebnis haben. Im Hypothesennetz wird daher ein deutlicher Einfluss des Laborbesuches auf das biologische Fähigkeitsselbstkonzept und die experimentbezogene Selbstwirksamkeit postuliert.

Bekannt ist, dass das Interessenprofil ein guter Prädiktor für die spätere Berufswahl ist (Ludwig 2003). Folglich wird auch ein deutlicher Zusammenhang des aktuellen Interesses und der durch den Laborbesuch bedingten Einstellungsänderung zu naturwissenschaftlichen Berufen angenommen. Wie dem Eccles-Modell (Eccles et al. 1983, vgl. Punkt 1.2.3.3) zu entnehmen ist, wird auch der Selbsteinschätzung ein Einfluss auf Leistungsmotivation und das Wahlverhalten (Entscheidung, sich mit bestimmten Aufgaben auseinanderzusetzen) zugeschrieben. Aus diesem Grund wird ebenfalls ein Einfluss der experimentbezogenen Selbstwirksamkeitserwartung sowie des Selbstkonzepts auf die Entscheidung für naturwissenschaftliche Berufe angenommen.

Das hier aufgestellte Hypothesennetz führt zu der Fragestellung, wie gut es die im BeLL Bio erhobenen Daten affektiver Einflussgrößen repräsentiert.

Von weiterem Interesse in der vorliegenden Studie ist das Potential von Schülerlaborkursen, ein heterogenes Publikum gleichermaßen fördern zu können. In Abschnitt 1.2.3.1 wurden Schülerlabore als Möglichkeit aufgezeigt, um Schüler unabhängig von Geschlecht und Migrationshintergrund zu erreichen. Insbesondere der Berücksichtigung des Migrationshintergrundes wurde in bisherigen Schülerlaboruntersuchungen nur relativ wenig Beachtung geschenkt. Neben den Faktoren Geschlecht und Migrationshintergrund wird in der vorliegenden Studie auch der naturwissenschaftlichen oder nicht naturwissenschaftlichen Berufsorientierung der Schüler Beachtung geschenkt, so dass sich als weitere Fragestellung ergibt:

Gelingt es den BeLL Bio-Kursen, ein heterogenes Schülerfeld unabhängig von Geschlecht, Migrationshintergrund und beruflicher Orientierung gleichermaßen anzusprechen?

Eine letzte Fragestellung ergibt sich aus der Beobachtung, dass Demonstrationsversuche im naturwissenschaftlichen Unterricht häufiger sind als Schülerversuche. Insbesondere im

Biologieunterricht experimentieren Schüler nur selten selbst (Barth, Pfeiffer 2009). Neben der geringen Experimentierfähigkeit der Schüler wird auch die möglicherweise geringe Experimentierfähigkeit der Lehrkräfte als Ursache in Erwägung gezogen (Scharfenberg 2005). Einen Eindruck von der Experimentiersicherheit kann die experimentbezogene Selbstwirksamkeitserwartung der Lehrkräfte vermitteln. Interessant ist in diesem Zusammenhang die folgende Fragestellung:

Steht die experimentbezogene Selbstwirksamkeitserwartung in Zusammenhang mit der von diesen Lehrkräften präferierten Experimentierform sowie mit der Experimentierhäufigkeit?

Unter Punkt 2.2 wird das in dieser Studie gewählte methodische Vorgehen zur Beantwortung der oben hergeleiteten fachdidaktischen Fragestellungen erläutert.

2 Material und Methoden

Wie bereits der Einleitungsteil, so gliedert sich auch dieser Abschnitt in einen fachwissenschaftlichen und einen fachdidaktischen Abschnitt. Im fachwissenschaftlichen Teil werden Material und Methoden der im Rahmen der Laborkursvorbereitung notwendigen fachwissenschaftlichen Arbeiten vorgestellt. Anschließend wird im fachdidaktischen Abschnitt das methodische Vorgehen zur Evaluierung der Laborkurse präsentiert.

2.1 Fachwissenschaftliche Untersuchungen

In diesem Abschnitt werden die molekularbiologischen und anschließend die photometrischen Methoden der fachwissenschaftlichen Untersuchungen dargestellt.

2.1.1 Molekularbiologische und physiologische Methoden

Um geeignete DNA-Abschnitte zu finden, die einen Genetischen Fingerabdruck in einer heterogenen Population wie einer Schulklasse zulassen, war es zunächst notwendig, geeignete Sequenzbereiche im menschlichen Genom zu wählen. Es wurden verschiedene VNTR-Loci (*variable number tandem repeats*) auf diese Bedingung hin untersucht. Dazu waren die Auswahl und Untersuchung geeigneter Primer und die Optimierungen der PCR-Bedingungen ebenso notwendig, wie die Wahl einer ungefährlichen und trotzdem zuverlässigen Färbemethode für die DNA-Banden. Darüber hinaus musste eine Methode zur effektiven, ungefährlichen und kostengünstigen Isolierung von DNA gefunden werden. Zunächst werden daher die molekularbiologischen Methoden beschrieben.

2.1.1.1 Gewinnung des Probenmaterials zum Genetischen Fingerabdruck

Unterschiedliche ethnische Gruppierungen können sich in bestimmten VNTR-Loci in der Allelfrequenz unterscheiden (Das & Mastana 2003). Um sicher zu stellen, dass im Schülerlabor ein auswertbares Ergebnis erzielt werden kann, wurden die DNA-Proben von freiwilligen Spendern kaukasischer, afrikanischer und asiatischer Herkunft nach dem Zufallsprinzip entnommen, um eine mögliche Situation in einer Schulklasse zu simulieren.

Als Probenmaterial dienten Epithelzellen der humanen Buccalschleimhaut. Diese wurden durch eine Mundspülung mit 25 mL dest. H₂O (Dauer: 15 Sekunden) gewonnen und dann in ein Falconröhrchen (50 mL) überführt. Die Sedimentierung der Zellen erfolgte durch Zentrifugation bei 3.773 x g in einer

Zentrifuge (Fa. Eppendorf) mit anschließender Verwerfung des Überstandes. Das Pellet wurde für Versuche zur DNA-Isolierung eingesetzt.

2.1.1.2 Methoden der DNA-Extraktion

2.1.1.2.1 DNA-Extraktion mit dem DNeasy Plant Mini Kit

Die DNA der Mundschleimhautzellen wurde mit dem DNeasy Plant Kit (Fa. Qiagen) isoliert. Die genaue Zusammensetzung der Lösungen ist nicht bekannt. Das Pellet wurde in flüssigem Stickstoff kurz schockgefroren und anschließend durch Mörsern aufgebrochen. Bei der nachfolgenden Zelllyse werden die Zellen mittels eines Detergens (Buffer AP1) aufgeschlossen. Anschließend erfolgt der enzymatische Abbau aller zellulären RNAs durch eine RNase (4 µL RNase A 100 mg/mL). Alle Proteine werden im folgenden Schritt gefällt (Buffer AP2) und durch Zentrifugieren über eine im Kit befindliche "QIAshredder-Säule" zurückgehalten. Nachfolgend wird die DNA mittels Ionenaustausch-Chromatographie unter hohen Salzkonzentrationen (Buffer AP3/E) an die Silikagel-Membran der "QIAshredder Spin-Säule" gebunden. Durch mehrere Waschschriffe werden Reste von Proteinen, Polysacchariden und Salzen entfernt. Die DNA wird mit Puffer (AE) eluiert, der eine geringe Salzkonzentration aufweist. Um eine höhere DNA Ausbeute zu erhalten, wurde der Elutionsschritt einmal wiederholt. Die erhaltenen Eluate werden zu je 50 µL aliquotiert und bei -20°C bis zur weiteren Verwendung gelagert.

2.1.1.2.2 DNA-Extraktion mit Chelex

Zur Freisetzung der DNA aus den Epithelzellen der Mundschleimhaut wurde das Zellpellet in 500 µL einer 5%igen Chelex®100-Lösung (Fa. Biorad, München) aufgenommen. Als Komplexbildner bindet Chelex in der Probe vorhandene polyvalente Kationen, welche später die PCR stören würden. Anschließend wurden 5µL Proteinase K (20mg/mL ddH₂O) zugesetzt. Der Ansatz wurde kurz gevortext und für 3 h bei 56 °C im vorgeheizten Thermoblock (Fa. Biometra) inkubiert. Nach erneutem kurzen Vortexen erfolgte die Denaturierung der Proteinase K für 10 Min. in einem auf 100 °C vorgeheizten Thermoblock (Fa. Biometra). Im Anschluss wurden die Proben kurz gevortext und für 5 Min. bei 9.391 x g in einer Zentrifuge (Fa. Eppendorf) abzentrifugiert. Das Chelex setzt sich als Pellet ab. Vom DNA-haltigen Überstand wurde vorsichtig – ohne Aufwirbelung der Chelexharz-Kügelchen – 300 µL abpipettiert, aliquotiert und bis zum Einsatz bei -20°C gelagert.

2.1.1.2.3 DNA-Extraktion durch Kaliumacetat-Isopropanol-Fällung

Diese traditionellere, günstige und gefahrlose Methode wurde optimiert, um ein im Schülerlabor einsetzbares Verfahren zu finden. Im Laufe der Optimierungen wurden zahlreiche Variationen mit der Menge des Zellextraktes, Inkubationszeiten und Bedingungen des Zentrifugierens untersucht. Hier wird die finale, als optimal betrachtete Methode dargestellt. Die zusätzlichen Versuchsreihen finden sich in Tab. 6.

25 mL Speichelprobe werden bei 3.773 x g 5 Min. zentrifugiert (Fa. Eppendorf) und der Überstand dekantiert. Die Zellen werden mit 500 µL Lysepuffer L (50 mM Tris, 10mM EDTA, 2% SDS, eingestellt auf pH8 mit HCl) lysiert. Die Proteine dieser Lösung (500 µL) werden mit 100 µL Fällungspuffer (8 M Kaliumacetat) auf Eis (5 Min.) gefällt. Reaktionsgefäße werden 5 Min. bei 9.391 x g zentrifugiert, 500 µL Überstand abgezogen, in ein neues Reaktionsgefäß überführt und mit 300 µL Isopropanol zur Fällung der DNA versetzt. Die Lösung wird 10 Minuten bei 9.391 x g zentrifugiert, der Überstand dekantiert und das Pellet mit 1 mL 70 % Ethanol gründlich gemischt. Anschließend wird 3 Minuten bei 9.391 x g zentrifugiert, der Überstand vorsichtig dekantiert und das Pellet – bei offenem Deckel – im Heizblock bei 50°C für 15 Minuten zum Trocknen belassen. Das Pellet wird in 100 µL autoklaviertem (von DNAsen befreiten), bidestilliertem Wasser gelöst und auf Eis gestellt.

Tabelle 6: Variationen bei der Isolierung von DNA aus Buccalschleimhautzellen

Variationen	1	2	3	4
Volumen Speichelprobe	40 mL	40 mL	30 mL	25 mL
1. Zentrifugation (Zellen pelletieren)	10 min. 3.773 x g	5 min. 3.773 x g	5 min. 3.773 x g	5 min. 3.773 x g
Proteinfällung auf Eis	10 min	10 min	5 min	5 min
2. Zentrifugation (Proteine sedimentieren)	10 min. 9.391 x g	10 min. 2.348 x g	10 min. 9.391 x g	5 min. 9.391 x g
3. Zentrifugation (DNA pelletieren)	15 min 9.391 x g	15 min 9.391 x g	10 min 9.391 x g	10 min 9.391 x g

2.1.1.3 PCR (polymerase chain reaction)

2.1.1.3.1 Primer als Startpunkte für die PCR

Mit Hilfe der von Mullis et al. (1986) entwickelten Polymerasen-Ketten-Reaktion (*polymerase chain reaction*, PCR) ist es möglich, beliebige DNA-Abschnitte innerhalb kürzester Zeit selektiv und exponentiell zu amplifizieren. PCR-Bedingungen, die in einem Labor ideal erscheinen, bedürfen ggf. unter den Gegebenheiten eines anderen Labors der Optimierung (vgl. Dekan et al. 1992b). Um optimale Bedingungen für den Einsatz im Schülerlabor herauszufinden, wurden die Bedingungen der

PCR in Abhängigkeit vom untersuchten Locus und der Primerzusammensetzung variiert. Je nach verwendetem DNA-Abschnitt kamen verschiedene Primer zum Einsatz (Tab. 7). Die Primer wurden von MWG Eurofins bezogen.

Tabelle 7: PCR-Primer

Name	Sequenz 5' - 3'	Fragmentgröße (ca.)	Referenz
<u>D17S30 VNTR</u>			
5'	AAA CTG CAG AGA GAA AGG TCG AAG AGT GA	168-1040 bp	Buscemi et al. (1994), mod.
3'	AAA GGA TCC CCC ACA TCC GCT CCC CAA GT		Buscemi et al. (1994), mod.
<u>apoB VNTR</u>			
5'	ATG GAA ACG GAG AAA TTA TG	570-900 bp	Boerwinkle et al. (1989)
3'	CCT TCT CAC TTG GCA AAT AC		Boerwinkle et al. (1989)
<u>Ig - JH VNTR</u>			
5'-Ig-J _H	GGG CCC TGT CTC AGC TGG GGA	470-920 bp	Decorte et al. (1990)
3'-Ig-J _H	TGG CCT GGC TGC CCT GAG CAG		Decorte et al. (1990)
<u>DIS80 VNTR</u>			
5'	GAA ACT GGC CTC CAA ACA CTG CCC GCC G	200-700 bp	Budowle et al. (1991)
3'	GTC TTG TTG GAG ATG CAC GTG CCC CTT GC		Budowle et al. (1991)

Die optimale Annealingtemperatur kann anhand der folgenden Formel berechnet werden:

$$T_m = [4^{\circ}\text{C} \times (\text{G} + \text{C}) + 2^{\circ}\text{C} \times (\text{A} + \text{T})] \quad (\text{vgl. Mülhardt 2009})$$

Allerdings erfordern einige Eigenschaften des Primers (Sequenzfolge etc.) sowie der eingesetzten DNA oftmals eine von der idealen Temperatur abweichende Bedingung (vgl. Mülhardt 2009).

2.1.1.3.2 Gradienten-PCR

Zur Ermittlung der idealen Annealingtemperatur wurde daher eine Gradienten-PCR (Fa. Eppendorf) durchgeführt. Diese erlaubt es, verschiedene Temperaturen in einem PCR-Lauf unter sonst identischen Bedingungen einzusetzen. In diesem Fall wurde aufgrund der Basenzusammensetzung der Primer ein Bereich von 45-65,5 °C Annealingtemperatur gewählt. Es wurden hier jeweils 5 µL DNA als Template mit folgendem PCR-Protokoll (Tab. 8) und hold bei 4 °C eingesetzt. Die Primer hatten eine Konzentration von 10 pmol/µL. Für diese und die weiteren PCR Reaktionen wurden neben den schon beschriebenen Primern weitere Substanzen benötigt. Diese wurden zur Vereinfachung der Arbeitsabläufe bei mehreren Proben in einem Mastermix zusammengefasst (Tab. 9). Der Mastermix enthielt alle für die PCR benötigten Substanzen, außer der DNA (diese wurde separat in die PCR Reaktionsgefäße pipettiert).

Tabelle 8: Protokoll Gradienten-PCR

Gradienten-PCR	Temp. [°C]	Zeit[min.]
Initialdenaturierung	95	5
Amplifizierung 34 x		
Denaturierung	94	1
Annealing	45,0; 50,4; 55,8; 61,0; 65,6	1
Elongation	72	1
Finale Elongation	72	10

**Tabelle 9: PCR Mastermix (Endvolumen 25 µL).
Die DNA wurde später zugegeben.**

Chemikalien	Einsatz [µL]
Primer (10 pmol/µL)	2,0
Puffer (10x)	2,5
MgCl ₂ (25 mM)	2,5
dNTP (20 pmol/µL)	1,0
ddH ₂ O	varierte
DNA	1-5
Taq-Polymerase 5 U/µL	1,0

2.1.1.3.3 PCR - Bedingungen

Nachdem für die gewählten Primerkombinationen geeignete Annealingtemperaturen ermittelt wurden, konnten die PCRs zum molekularen Identitätsnachweis begonnen werden. Dabei waren weitere Variationen der Anzahl der Zyklen, der Zeiten und Temperaturen notwendig (Tab. 10-13). Als DNA-Template wurden Volumina zwischen 1 und 5 µL DNA-Eluat der drei verschiedenen Extraktionsmethoden Qiagen Plant Mini Kit, Chelex-Methode und Kaliumacetat-Isopropanol-Fällung eingesetzt. Die Variationen der DNA-Volumina (Tab. 9) sind aus Gründen der besseren Lesbarkeit nicht zusätzlich bei den Variationen der PCR-Bedingungen dargestellt.

Tabelle 10: PCR-Programme für den D17S30 VNTR Locus. Die Spalten Temperatur und Zeit enthalten (getrennt durch /) entsprechende Variationen. Variierende Zyklenzahlen stehen unter Amplifizierung.

PCR (vgl. Buscemi et al. 1994)	Temp. [°C]	Zeit[min]
Initialdenaturierung	95	5
Amplifizierung 27 / 29 / 30x		
Denaturierung	94	1
Annealing	63 / 56 / 65	2 / 2 / 1
Elongation	72	2 / 2 / 1
Finale Elongation	63 / 72 / 72	11 / 10 / 10

Tabelle 11: PCR-Programme für den apoB VNTR Locus. Die Spalten Temperatur und Zeit enthalten (getrennt durch /) entsprechende Variationen.

PCR (vgl. Boerwinkle et al. 1989)	Temp. [°C]	Zeit[min]
Initialdenaturierung	95	7 / 5 / 5
Amplifizierung 30 x		
Denaturierung	95 / 94 / 94	1
Annealing	55 / 55 / 51	1
Elongation	72	2
Finale Elongation	72	7 / 10 / 10

Tabelle 12: PCR-Programme für den Ig - JH VNTR Locus. Die Spalten Temperatur und Zeit enthalten (getrennt durch /) entsprechende Variationen. Variierende Zyklenzahlen stehen unter Amplifizierung

PCR (vgl. Decorte et al. 1990)	Temp. [°C]	Zeit[min]
Initialdenaturierung	95	5
Amplifizierung 30 / 35 / 35 x		
Denaturierung	94	0,5 / 1
Annealing	63 / 56 / 65	1
Elongation	72	2 / 1
Finale Elongation	72	10

Tabelle 13: PCR-Programme für den D1S80 VNTR Locus. Die Spalten Temperatur und Zeit enthalten (getrennt durch /) entsprechende Variationen. Variierende Zyklenzahlen stehen unter Amplifizierung

PCR (vgl. Budowle et al. 1991)	Temp. [°C]	Zeit[min]
Initialdenaturierung	95	5
Amplifizierung 25x / 30x		
Denaturierung	95 / 94	1
Annealing	65	1
Elongation	70 / 72	8 / 2
Finale Elongation	72	10

2.1.1.4 Gelelektrophorese

Bei der Gelelektrophorese wandern Ionen bzw. geladene Moleküle bei angelegter Spannung durch ein definiertes Medium. Nukleinsäuren wandern in einem elektrischen Feld aufgrund der negativen Ladung ihrer Phosphatgruppe zur Anode. Mittels der Agarose-Gelelektrophorese (Sambrock et al. 1989) lassen sich DNA-Moleküle der Größe nach auftrennen. Agarose ist ein Polysaccharid, welches aus roten Meeresalgen gewonnen wird. Beim Aufkochen in Puffer (Puffer schafft später die optimale Umgebung für die DNA-Fragmente) verflüssigt es sich. Beim Abkühlen erstarrt es – in eine Form gegossen - zu einem großporigen Gel. Das erstarrte Gel kann anschließend an einem Ende mit DNA-Proben beladen werden. Die Laufstrecke der DNA-Moleküle richtet sich nach ihrer Länge sowie nach

ihrer Konformation. Doppelsträngige DNA-Moleküle werden in der Elektrophorese nur der Größe nach aufgetrennt. Sind allerdings ringförmig geschlossene Plasmide in der Probe enthalten, so verhalten sich diese in der Gelelektrophorese wie deutlich kleinere lineare Moleküle. Dies geschieht aufgrund ihrer kompakten "supercoiled"-Konformation (vgl. Ma & Bloomfield 1995).

2.1.1.4.1 Tris-Acetat-EDTA-Puffer (TAE-Puffer)

Folgende Puffer wurden für die Gelelektrophorese verwendet:

Tabelle 14: Zusammensetzung des TAE-Puffers (Laufpuffer), 50-faches Konzentrat

Konzentration	Zutat
243,3 g/L	Tris 40 mM
16,62 g/L	EDTA (Titriplex III) 1 mM

Der pH-Wert des Puffers wurde mit Eisessig auf pH 7,8 eingestellt, der Puffer autoklaviert und anschließend bei RT (Raumtemperatur) gelagert. Die Gelelektrophorese zur Trennung der DNA wurde in Midi 1 Elektrophorese-Kammern (Fa. Roth) mit Gelen von 10 x 11,5 cm (B x L) Größe und einem Kammervolumen von 450 mL durchgeführt. Für die Elektrophorese wurden Gele zwischen 1 und 2 % Agarose (w/v; Fa. Roth) hergestellt. Dazu wurde die Agarose in 1x Tris-Acetat-EDTA (TAE)-Puffer (Tab. 14) durch Aufkochen in der Mikrowelle gelöst. Für die Verwendung im BeLL Bio wurden verschiedenen Färbemethoden getestet.

2.1.1.4.2 Farbstoffe zur Sichtbarmachung der DNA

Als Farbstoffe zur Färbung der DNA-Banden wurden verschiedene, meist in die DNA interkalierende Substanzen eingesetzt (Tab. 15).

Tabelle 15: Überblick der getesteten Farbstoffe und deren Konzentrationen

Farbstoff	Konzentration
Ethidiumbromid	1 µL auf 100 mL Agarose
Stain G	0,5 µL auf 50 mL Agarose
Gel Red	5 µL auf 50 mL Agarose
Methylenblau	0,02 % (w/v)

Ethidiumbromid-Färbung: Nach Abkühlen der Agarose auf ca. 54 °C wurde Ethidiumbromidlösung (Fa. Roth, 1%ige Lösung, 1 µL zu je 100 mL Agarose) hinzugegeben. Bei Bedarf wurde der Laufpuffer (TAE-Puffer) ebenfalls mit Ethidiumbromid versetzt (1 µL je Kammerfüllung von 450 mL). Vorteil von Ethidiumbromid ist seine hohe Nachweisempfindlichkeit und sein günstiger Preis (10 ml 1%ige

Lösung ca. 8 €). Da Ethidiumbromid aufgrund seiner interkalierenden Fähigkeiten allerdings sehr giftig ist und vermutlich genetische Defekte verursachen kann, wurden alternative Färbemethoden für den Einsatz im BeLL Bio getestet.

Gel Red-Färbung: Nach Abkühlen der Agarose auf ca. 54 °C wurde sie mit Gel-Red-Lösung (Fa. Biotium, 10.000x in H₂O, 10 µL je 100 mL Agarose) versetzt. Gel Red ist deutlich teurer als Ethidiumbromid (0,5 mL 10.000x kosten ca. 170€), besitzt eine sehr hohe Nachweisempfindlichkeit und gilt im Gegensatz zu Ethidiumbromid als nicht zellmembrandurchlässig und nicht cytotoxisch (www.biotium.com).

Stain G-Färbung: Nach Abkühlen der Agarose auf ca. 54 °C wurde diese mit Stain G-Lösung (Fa. Serva, 1 µL je 100 mL Agarose) versetzt und anschließend das Gel gegossen. Ähnlich dem Ethidiumbromidverfahren wurde Stain-G Lösung (1 µL je Kammerfüllung von 450 mL) direkt vor dem Anlegen der Spannung in den Laufpuffer (TAE-Puffer) der Kammer gegeben (auf die Seite der Katode). Stain G besitzt wie die zuvor genannten Farbstoffe eine hohe Nachweisempfindlichkeit, gilt als nicht karzinogen (www.serva.de) und liegt preislich zwischen den zuvor genannten Alternativen. Da die Färbung mit Stain G sehr gute Ergebnisse lieferte, wurden alle weiteren Versuche mit Stain G zur Sichtbarmachung der DNA-Banden durchgeführt.

Methylenblau-Färbung: Nach Abkühlen der Agarose auf ca. 54 °C wurde das Gel gegossen, und die Elektrophorese durchgeführt. Erst anschließend wurde das Gel für 30 min. in der vorbereiteten Methylenblaulösung (0,02 %) gefärbt, bevor es für weitere 30 min. in Wasser wieder soweit entfärbt wurde, dass die DNA-Banden im Gel sichtbar wurden (Wasserwechsel nach 15 Minuten).

Als Ladepuffer diente – unabhängig von der Färbemethode – der 6x Orange Loading Dye (Fa. Fermentas) oder der 6x Loading Dye (Fa. Fermentas). Je 1 µL Ladepuffer wurde zu 5 µL Probe aus der PCR gegeben, gemischt und dann in die Taschen des Agarosegels gegeben. Anschließend wurde die DNA für ca. 60 min bei 100 V, 150W und variierender Stromstärke (A) aufgetrennt. Als Laufpuffer diente 1xTAE-Puffer (Tab. 14).

Nach erfolgter Gelelektrophorese werden die in die DNA interkalierenden Farbstoffe (Methylenblau interkaliert nicht und ist als blauer Farbstoff direkt sichtbar) mittels UV-Durchlicht Apparatur detektiert ($\lambda=312$ nm) und mit der angeschlossenen CCD-Kamera (Fa. Intas) dokumentiert.

2.1.1.4.3 Marker zur Größenbestimmung

Zusätzlich zu den zu untersuchenden DNA-Proben wurden verschiedene DNA-Marker (Ladder) vor der Gelelektrophorese mit auf das Gel aufgetragen. Diese enthalten DNA-Fragmente definierter Größe und dienen als Längenvergleich für die eigentlichen PCR-Produkte. Je nach erwarteter Bandengröße wurden verschiedene DNA-Marker (Ladder) als Längenstandard verwendet:

- Fragmentgröße 250 – 10.000 bp: 3 µL GeneRuler™ 1 kb DNA Ladder Mix (MBI Fermentas; 0,5 mg DNA/mL)
- Kleinere DNA-Fragmente: 3 µL GeneRuler™ DNA Ladder Low Range (MBI Fermentas; 0,5 mg DNA/mL)
- 100 bp DNA-Leiter äquimolar (Fa. Roth; 0,1 mg DNA/mL).

Die zuletzt genannte 100 bp-Ladder wurde darüber hinaus eingesetzt, um die Empfindlichkeit der oben genannten DNA-Farbstoffe zu testen. Dazu wurden verschiedene Mengen der Leiter (0,05 µg DNA, 0,1 µg DNA, 0,2 µg DNA, 0,3 µg, 0,4 µg und 0,5 µg) auf ein mit dem entsprechenden UV-Farbstoff versetztes Gel pipettiert. Nach der Gelelektrophorese wurde das Bandenmuster dokumentiert. Entsprechendes Vorgehen wurde mit jedem der oben genannten Farbstoffe durchgeführt. Lediglich die Färbung mit Methylenblau erfolgte nach der Elektrophorese.

2.1.2 Photometrische Messungen zum Pyruvatabbau durch die Laktatdehydrogenase

Für den Laborkurs „Außer Atem – Sportliche Enzyme“ (LDH) wurde eine zeitökonomische Möglichkeit gesucht, um die im Kurs untersuchten Pyruvat-Proben rot anzufärben. Die so gefärbte Pyruvat-Lösung sollte den Eindruck verdünnter Blutproben vermitteln, um das Interesse der Schüler anzuheben und die Authentizität zu erhöhen. Die Identifizierung geeigneter Farbstoffe erfolgte photometrisch. Zur Anfärbung der Pyruvatlösung wurden keine Reinstoffe verwendet, sondern intensiv rot gefärbte, ungiftige Erfrischungsgetränke. Tabelle 16 gibt einen Überblick über die verwendeten Materialien.

Tabelle 16: Material für die photometrischen Versuche

Material
Photometer Genesys 10S UV/Vis (Fa. Thermo Scientific)
Einmal-Küvetten (PMMA, Durchlässigkeit: 280-800 nm)
Erfrischungsgetränk Powerade Zero (enthält Lebensmittelfarbstoff E163, Anthocyane, unverdünnt)
Erfrischungsgetränk Powerade Wild Cherry (enthält Lebensmittelfarbstoff E122, Azorubin, 1:4-Verdünnung)
Erfrischungsgetränk Punica Classics Rote Früchte (Farbstoffe nicht exakt definiert, unverdünnt)
L-lactic Dehydrogenase from bovine muscle, 5,5 mg/mL, ≥ 600 U/mg Protein (Fa. Sigma-Aldrich), verdünnt auf ~ 3 U/50 μ L
NADH- Na_2 0,013 mol/L
Imidazolpuffer, pH 7,5, 1 mol/L (pH eingestellt mit HCl)
Natriumpyruvat
ddH ₂ O

Zunächst sollten Farbstoffe mit einer zu starken Absorption im gewünschten Wellenlängenbereich der NADH-Messung identifiziert und für die weitere Verwendung ausgeschlossen werden. Dazu wurden Absorptionsspektren der in Tab. 16 genannten Erfrischungsgetränke Powerade Wild Cherry (1:4-Verdünnung), Powerade Zero (unverdünnt) und Punica Classics (unverdünnt) im Wellenlängenbereich von $\lambda=300-800$ nm gegen einen Blindwert aus ddH₂O vermessen.

Weiterhin galt es auszuschließen, dass die zur Färbung als geeignet identifizierten Getränke zu einer unvollständigen Umsetzung des Pyruvats führen. In diesem Fall wäre eine Berechnung der ursprünglich in der Probe vorhandenen Pyruvatkonzentration anhand des photometrisch beobachteten NADH-Verbrauches nicht möglich. Zu diesem Zweck wurden drei Natriumpyruvatproben mit unterschiedlichen Lösungsmitteln angesetzt. Für Probe A wurde Pyruvat in ddH₂O gelöst. Die Konzentration der Pyruvat-Probe A betrug gemäß Einwaage $c_{A_Einwaage}=0,014$ mol/L. Probe B wurde in einer 1:4-Verdünnung von Powerade Wild Cherry angesetzt ($c_{B_Einwaage}=0,014$ mol/L). Pyruvat der Probe C wurde in Powerade Zero gelöst ($c_{C_Einwaage}=0,018$ mol/L). Die Bestimmung der Pyruvatkonzentration erfolgte anschließend indirekt über die photometrische Bestimmung des NADH-Verbrauchs beim Umsatz von Pyruvat zu Laktat durch die Laktatdehydrogenase. Die Zusammensetzung der Ansätze und der entsprechenden Blindwerte für die photometrische Messung ergibt sich aus Tabelle 17. Sämtliche Messungen erfolgten bei Raumtemperatur. NADH und LDH wurden zwischen den Messungen bei 4 °C gelagert.

Tabelle 17: Zusammensetzung der Ansätze für die photometrische Messung des NADH-Verbrauchs

	Blindwert A (ohne Powerade)	Ansatz A (ohne Powerade)	Blindwert B (Powerade Wild Cherry)	Ansatz B (Powerade Wild Cherry)	Blindwert C (Powerade Zero)	Ansatz C (Powerade Zero)
ddH ₂ O	1850 µL	1725 µL	1850 µL	1725 µL	1850 µL	1725 µL
Imidazolpuffer (1 mol/L, pH 7,5)	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL
NADH, 0,013 mol/L	---	75 µL	---	75 µL	---	75 µL
Pyruvat in dH ₂ O gelöst	50 µL	50 µL	---	---	---	---
Pyruvat in Powerade Wild Cherry gelöst	---	---	50 µL	50 µL	---	---
Pyruvat in Powerade Zero gelöst	---	---	---	---	50 µL	50 µL
LDH (Zusatz nach t=1 min)	---	50 µL	---	50 µL	---	50 µL

Zunächst wurde photometrisch bei einer Wellenlänge von $\lambda=366$ nm die NADH-Absorption im Ansatz ohne Anwesenheit von LDH bestimmt. Nach t=1 min wurden dem Ansatz 50 µL der LDH-Lösung zugesetzt. Der NADH-Abbau wurde so lange verfolgt, bis kein Absinken der Absorption mehr beobachtet werden konnte. Die Messung für jede Pyruvatprobe wurde fünfmal durchgeführt und die durchschnittliche Absorptionsänderung ΔAbs_0 in das Lambert-Beersche-Gesetz eingesetzt. Die berechnete Konzentrationsänderung des NADH wurde verwendet, um auf die Ausgangskonzentration an Pyruvat in der Probe zu schließen. Durch Abgleich der photometrisch bestimmten ($c_{\text{photometrisch}}$) mit der per Einwaage berechneten Pyruvatkonzentration (c_{Einwaage}), konnte in jedem der drei Ansätze festgestellt werden, ob sämtliches Pyruvat in der Probe zu Laktat umgesetzt wurde.

2.2 Fachdidaktische Untersuchungen

Im fachdidaktischen Methodenabschnitt werden zunächst die Stichprobe, die Messinstrumente, das Versuchsdesign und die drei untersuchten Schülerlaborkurse vorgestellt. Abschließend werden die Verfahren der Datenauswertung beschrieben.

2.2.1 Beschreibung der Stichprobenzusammensetzung

Die vorliegende Untersuchung setzt sich aus einer Haupt- und einer Nebenstudie zusammen. In beiden Studien wurden Schüler untersucht, welche im Rahmen ihres Unterrichts das Schülerlabor Bell Bio an der Bergischen Universität Wuppertal besuchten. In Anlehnung an vergleichbare Studien (Brandt 2005, Scharfenberg 2005) entspricht das hier berichtete Vorgehen einer quasiexperimentellen Untersuchung. Dies bedeutet, dass die Schüler nicht einer rein zufällig

ausgewählten Stichprobe angehören, sondern dass – im Sinne einer Klumpenstichprobe (Bortz & Döring 2006) – „natürlich gewachsene“ Schulklassen untersucht werden. In Ergänzung zu den beiden Schülerbefragungen wurde eine Fragebogenerhebung zur Experimentiersicherheit von Lehrkräften durchgeführt.

2.2.1.1 Stichprobe der Hauptstudie

An der Hauptstudie nahmen 842 Schüler aus 38 Biologiekursen der Oberstufe nordrhein-westfälischer Schulen teil. Aufgrund der Datenerhebung über drei Messzeitpunkte ergab sich eine Stichprobenmortalität vom ersten zum zweiten Messzeitpunkt von 25,06%. Die 631 Personen ($N_{\text{weiblich}}=361$, $N_{\text{männlich}}=270$, Durchschnittsalter=17 Jahre), die zu den ersten beiden Messzeitpunkten befragt werden konnten, verblieben in der finalen Auswertung. Davon gehörten 520 Schüler der Oberstufe des Gymnasiums, 81 der gymnasialen Oberstufe an Gesamtschulen, 39 der Oberstufe am Berufskolleg (Berufliches Gymnasium) und 12 der Oberstufe einer Waldorfschule an. 132 dieser Schüler bildeten die Kontrollgruppe, während die übrigen sich auf die drei Treatmentgruppen Genetischer Fingerabdruck, Laktatdehydrogenase (LDH) und Einzeller verteilten. Die folgenden Tabellen geben einen Überblick über die genaue Zusammensetzung der Versuchsgruppen und der Kontrollgruppe in Bezug auf das Geschlechterverhältnis (Tab. 18) die besuchte Schulform (Tab. 19) und Jahrgangsstufe (Tab. 20).

Tabelle 18: Anzahl (N) der weiblichen und männlichen Schüler in den Treatmentgruppen und der Kontrollgruppe

Geschlecht	Treatment Gen. Fingerabdruck (N)	Treatment Einzeller (N)	Treatment LDH (N)	Kontrollgruppe (N)	Gesamt (N)
weiblich	128	41	107	85	361
männlich	96	34	93	47	270
Gesamt	224	75	200	132	631

Tabelle 19: Anzahl (N) der Schüler in den Treatmentgruppen und der Kontrollgruppe in Abhängigkeit von der besuchten Schulform

Schulform	Treatment Gen. Fingerabdruck (N)	Treatment Einzeller (N)	Treatment LDH (N)	Kontrollgruppe (N)	Gesamt (N)
Berufskolleg	21	0	18	0	39
Gymnasium	149	48	170	132	499
Gesamtschule	54	27	0	0	81
Waldorfschule	0	0	12	0	12
Gesamt	224	75	200	132	631

Tabelle 20: Anzahl (N) der Schüler in den Treatmentgruppen und der Kontrollgruppe in Abhängigkeit von der besuchten Jahrgangsstufe (EF=Einführungsphase, Q1=Qualifikationsphase 1 des achtjährigen Gymnasiums)

Jahrgangsstufe	Treatment Gen. Fingerabdruck (N)	Treatment Einzeller (N)	Treatment LDH (N)	Kontrollgruppe (N)	Gesamt (N)
11+EF	0	75	182	47	304
12+Q1	203	0	18	57	278
13	21	0	0	28	49
Gesamt	224	75	200	132	631

Bedingt durch die unterschiedlichen Zielgruppen der evaluierten Laborkurse gehören die Schüler in der Treatmentgruppe Genetischer Fingerabdruck überwiegend einer höheren Jahrgangsstufe an als diejenigen der Versuchsgruppen Einzeller und LDH (Tab. 20). Aktuell finden sich in der Oberstufe des Gymnasiums sowohl Schüler, die bereits das verkürzte Abitur (G8) absolvieren als auch parallel diejenigen, die nach neun Jahren das Gymnasium verlassen. Da die Einführungsphase (EF) nach G8 formal der früheren Jahrgangsstufe 11 und die Qualifikationsphase 1 (Q1) der früheren Jahrgangsstufe 12 entspricht, wird zwischen den entsprechenden Schülern in der Auswertung nicht differenziert. Das Verhältnis von Leistungs- zu Grundkursbesuchern gibt Tabelle 21 für die Schüler an, die im Untersuchungszeitraum bereits die Jahrgangsstufe 12, 13 oder Q1 besuchten.

Tabelle 21: Anzahl (N) der Leistungs- und Grundkursbesuchern in den Treatmentgruppen und der Kontrollgruppe (LK=Leistungskurs, GK=Grundkurs). Zusätzlich sind die Schüler jüngerer Jahrgangsstufen vermerkt (11+EF), die noch nicht in LK und GK differenziert werden können

Leistungs-/Grundkurs	Treatment Gen. Fingerabdruck (N)	Treatment Einzeller (N)	Treatment LDH (N)	Kontrollgruppe (N)	Gesamt (N)
LK	134	0	18	85	237
GK	90	0	0	0	90
11+EF	0	75	182	47	304
Gesamt	224	75	200	132	631

Zum dritten Messzeitpunkt reduzierte sich die Stichprobe auf 502 Personen. In längsschnittliche Untersuchungen über alle drei Messzeitpunkte fließt daher eine entsprechend reduzierte Personenzahl ein. Diese wird in den entsprechenden Tabellen im Ergebnisteil angegeben. Erst zum dritten Messzeitpunkt wurde der Migrationshintergrund der Schüler erfragt. Da in der Treatmentgruppe Einzeller nur Daten zu den ersten beiden Messzeitpunkten erhoben wurden, liegen für diese Gruppe keine Angaben zum Migrationshintergrund vor. Insgesamt machten 407 Schüler Angaben zu ihrem Migrationshintergrund. Wie sich diese auf die Treatmentgruppen und die Kontrollgruppe verteilen, wird in Tabelle 22 dargestellt.

Tabelle 22: Anzahl (N) der Schüler mit und ohne Migrationshintergrund in den Treatmentgruppen Genetischer Fingerabdruck und LDH sowie in der Kontrollgruppe

	Treatment Gen. Fingerabdruck (N)	Treatment LDH (N)	Kontrollgruppe (N)	Gesamt (N)
Migrationshintergrund	42	35	21	98
ohne Migrationshintergrund	129	95	85	309
Gesamt	171	130	106	407

Weiterhin wurde erhoben, ob die Schüler eine berufliche Laufbahn im naturwissenschaftlichen Bereich anstreben oder eher nicht. Hierzu liegen Angaben von 372 Schülern vor. Die Verteilung der Schüler auf die Treatmentgruppen und die Kontrollgruppe ist Tabelle 23 zu entnehmen.

Tabelle 23: Anzahl (N) der Schüler mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung in den Treatmentgruppen Genetischer Fingerabdruck und LDH sowie in der Kontrollgruppe

Berufliche Orientierung	Treatment Gen. Fingerabdruck (N)	Treatment LDH (N)	Kontrollgruppe (N)	Gesamt (N)
naturwissenschaftlich	40	33	31	104
nicht naturwissenschaftlich	115	81	72	268
Gesamt	155	114	103	372

2.2.1.1.1 Stichprobe der Videostudie von Dahmen (2012)

In die Datenauswertung der Hauptstudie wurden Daten einbezogen, welche von Dahmen (2012) im Rahmen einer Videoanalyse erhoben wurden. Die Videoaufzeichnungen wurden parallel zu den Fragebogenerhebungen der Hauptstudie in acht Laborkursen der Treatmentgruppe Genetischer Fingerabdruck vorgenommen. Gefilmt wurden alle Schüler dieser Kurse bei der Durchführung einer Pipettierübung sowie beim Beladen eines Agarose-Gels mit DNA-Proben. Ausgewertet wurden von Dahmen (2012) N=106 Videos für die Pipettierübung sowie N=101 Videos beim Beladen des Gels.

Tabelle 24: Stichprobenzusammensetzung der im Rahmen der Pipettierübung aufgezeichneten und den Fragebögen der Hauptstudie zugeordneten Videodaten (LK=Leistungskurs, GK=Grundkurs)

Geschlecht	Berufskolleg (N)	Gymnasium (N)	Gesamtschule (N)	Gesamt (N)
weiblich	10	33	11	54
männlich	10	18	11	39
Gesamt	20	51	22	93
Kursart	Berufskolleg (N)	Gymnasium (N)	Gesamtschule (N)	Gesamt (N)
LK	20	24	9	53
GK	0	27	13	40
Gesamt	20	51	22	93

Durch Nummerierung der Laborkittel in Kombination mit einer im Laborkurs geführten Liste, in welcher Kittelnummer und Geburtsdatum vermerkt waren, konnten N=93 Videos der Pipettierübung und N=89 Videos der Gelstation eindeutig den Fragebogendaten der Hauptstudie zugeordnet werden. Die Tabellen 24 und 25 geben einen Überblick über die Stichprobenzusammensetzung der in der vorliegenden Studie berücksichtigten Videodaten.

Tabelle 25: Stichprobenzusammensetzung der im Rahmen der Gelstation aufgezeichneten und den Fragebögen der Hauptstudie zugeordneten Videodaten (LK=Leistungskurs, GK=Grundkurs)

Geschlecht	Berufskolleg (N)	Gymnasium (N)	Gesamtschule (N)	Gesamt (N)
weiblich	10	36	11	57
männlich	6	15	11	32
Gesamt	16	51	22	89
Kursart	Berufskolleg (N)	Gymnasium (N)	Gesamtschule (N)	Gesamt (N)
LK	16	18	7	41
GK	0	33	15	48
Gesamt	16	51	22	89

2.2.1.2 Stichprobe der Nebenstudie

Die Nebenstudie wurde – ebenfalls als quasiexperimentelle Studie über drei Messzeitpunkte – ausschließlich im Laborkurs Genetischer Fingerabdruck durchgeführt. Die hier eingesetzten Messinstrumente erweiterten das Spektrum der Hauptstudie um die Untersuchung des kurzfristigen Einflusses des Laborbesuches auf die experimentbezogene Selbstwirksamkeit. An ihr nahmen 52 Gymnasialschüler ($N_{\text{weiblich}}=34$, $N_{\text{männlich}}=18$, Durchschnittsalter=17 Jahre) teil, die allesamt Biologieleistungskurse in der zwölften Jahrgangsstufe ($N=23$) beziehungsweise der Q1 ($N=28$) besuchten. Die Probanden waren zu allen drei Messzeitpunkten anwesend, so dass sich die Stichprobe nicht weiter reduzierte. Angaben über Migrationshintergrund und berufliche Orientierung wurden bei diesen Schülern nicht erhoben. Als Vergleichsgruppe dient in der Nebenstudie ebenfalls die bereits unter Punkt 2.2.1.1 beschriebene Kontrollgruppe.

2.2.1.3 Stichprobe der Lehrkräfte-Befragung

Die Stichprobe der Lehrer ($N=45$, Durchschnittsalter=41) setzt sich aus den Begleitlehrern der das BeLL Bio besuchenden Schulklassen ($N=27$) sowie aus Referendaren des Unterrichtsfachs Biologie ($N=18$) zusammen. Mit 30 weiblichen und 15 männlichen Lehrern ist ein erhöhter Frauenanteil gegeben. Vertreter von Gymnasien ($N=20$), Gesamt- ($N=10$), Real- ($N=8$) Haupt- ($N=5$) und Berufsbildenden Schulen ($N=2$) nahmen an der Erhebung teil. Da die Befragung nur zu einem Messzeitpunkt stattfand, ist keine Stichprobenmortalität gegeben.

2.2.2 Versuchsdesign der Evaluierungen

Nachdem im vorangegangenen Abschnitt die Stichproben der drei Teilstudien beschrieben wurden, wird nun das Versuchsdesign von Hauptstudie, Nebenstudie und der Lehrerbefragung erläutert.

2.2.2.1 Versuchsdesign der Hauptstudie

Die Hauptstudie wurde im Pre-, Post-, Follow-up-Test-Design angelegt, wobei zwischen Pre- und Posttest der Schülerlaborbesuch stattfand. Während die drei Treatmentgruppen (=Versuchsgruppen) Genetischer Fingerabdruck, Laktatdehydrogenase (LDH) und Einzeller den entsprechenden Laborkurs im BeLL Bio besuchten, wurde der Kontrollgruppe dieser Besuch vorenthalten (Abb. 23).

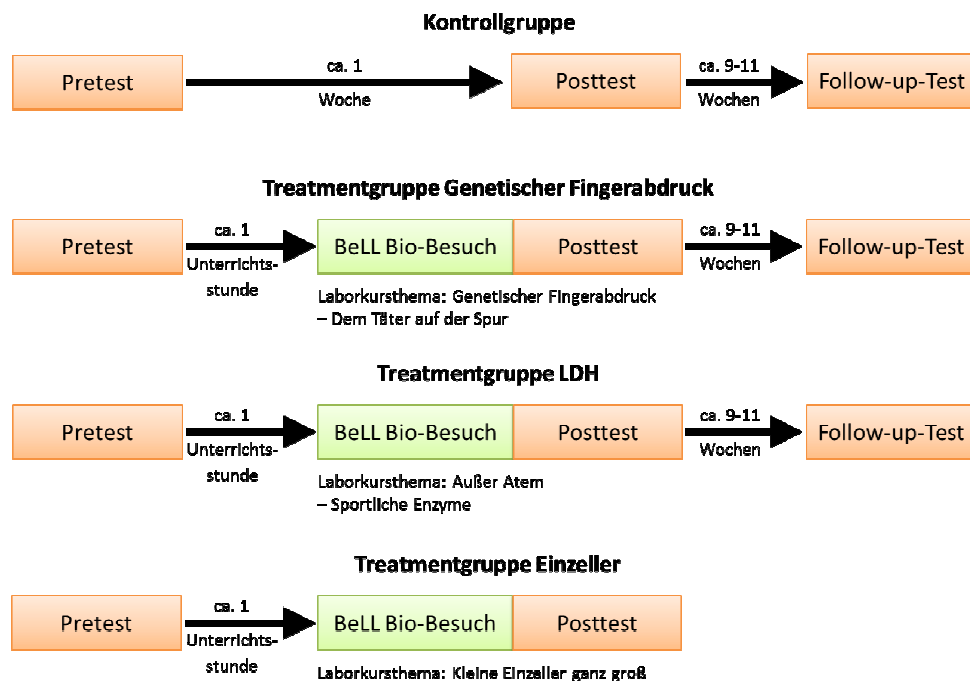


Abbildung 23: Versuchsdesign Hauptstudie

In den Treatmentgruppen wurde in der Regel in der vorletzten Biologiestunde vor dem Laborbesuch ein Pretest im Unterrichtsraum durchgeführt. Dieser erfasste über affektive Messinstrumente die persönliche Disposition und über einen kognitiven Wissenstest das individuelle Vorwissen der Schüler zu den Laborkursinhalten. Erst nach Durchführung des Pretests erhielten die Schüler ein laborkursspezifisches Versuchsskript. Die Versuchsanleitung wurde in der Unterrichtsstunde unmittelbar vor dem Laborbesuch mit der Lehrperson durchgesprochen. Im Anschluss daran folgte einmalig ein Besuch im Schülerlabor BeLL Bio an der Bergischen Universität Wuppertal. Die drei Treatmentgruppen unterschieden sich im Inhalt des jeweiligen Laborkurses. Nähere Informationen zu den Laborkursen finden sich unter Punkt 2.2.4. Unmittelbar im Anschluss an den Laborkurs wurde von den Schülern der Treatmentgruppen im Seminarraum des Schülerlabors ein Posttest ausgefüllt.

Dieser beinhaltete neben den Wissensfragen des Pretests die weiteren in Abschnitt 2.2.3.2.3 vorgestellten affektiven Messinstrumente. Im Abstand von etwa 9-11 Wochen nach dem Laborbesuch³ folgte ein Follow-up-Test, welcher langfristige kognitive und affektive Effekte des Laborbesuches erfasste. Einzige Ausnahme bildet hier die Treatmentgruppe Einzeller, in der kein Follow-up-Test durchgeführt wurde.

Bei der Kontrollgruppe wurde in gleicher Weise verfahren wie bei den Treatmentgruppen Genetischer Fingerabdruck und LDH, so dass auch hier zu den drei Messzeitpunkten jeweils Pre-, Post- und Follow-up-Tests durchgeführt wurden. Bei dieser Gruppe blieb der Laborbesuch allerdings aus. Die Kontrollgruppe hat zwei Funktionen. Sie dient einerseits als Vergleichsgruppe um feststellen zu können, ob beobachtete Längsschnitteffekte tatsächlich auf den Schülerlaborbesuch oder aber auf generelle zeitliche Schwankungen in der Ausprägung der beobachteten Merkmale zurückzuführen sind. Auch Pretesteffekte, wie beispielsweise ein durch das reine Ausfüllen des Wissens-Pretests bedingter Lerneffekt, können so kontrolliert werden. Darüber hinaus bietet die Kontrollgruppe die Möglichkeit, die Wahrnehmung der Lernumgebung Schülerlabor durch die Treatmentgruppen in Relation zu setzen zur Wahrnehmung der Lernumgebung schulischer Biologieunterricht.

2.2.2.1.1 Versuchsdesign Videostudie

In der vorliegenden Studie wird auf Daten der gemeinsam mit Dahmen (2012) durchgeführten BeLL Bio-Videostudie zurückgegriffen (vgl. Stichprobenbeschreibung unter Punkt 2.2.1.1.1). Die parallel zur Hauptstudie erhobenen Videodaten bieten die Möglichkeit, die in den Fragebögen erfasste experimentbezogene Selbstwirksamkeit mit der tatsächlich beobachteten Experimentierkompetenz der Schüler abzugleichen. Im Unterschied zur qualitativ ausgerichteten erziehungswissenschaftlichen Videographie (Dinkelaker & Herrle 2009) wurde ein quantitatives Design gewählt. Möglichst viele Schüler wurden dazu bei der Ausführung identischer Experimentierschritte im Laborkurs Genetischer Fingerabdruck gefilmt, um quantitative Aussagen über ihr Experimentierverhalten treffen zu können. Daher war es auch notwendig, solche Experimentierschritte auszuwählen, die von nahezu allen Schülern tatsächlich durchgeführt wurden. Hinsichtlich der räumlichen Gegebenheiten war zu berücksichtigen, dass der Laborraum, bedingt durch hohe Aufbauten der Laborbänke, nur sehr schlecht einseh- und filmbar war. Bei gleichzeitiger Experimentiertätigkeit an verschiedenen Arbeitsplätzen hätten diese daher nicht dokumentiert werden können. Deshalb wurden für die Videodokumentation zwei Experimentierschritte des Kurses Genetischer Fingerabdruck gewählt, die

³ Lediglich in zwei Kursen der Treatmentgruppe LDH musste der Follow-up-Test – bedingt durch die sich anschließenden Sommerferien – bereits einen Monat nach dem Laborbesuch durchgeführt werden. Aufgrund der neuen Kurszusammensetzung nach den Ferien wäre eine Befragung ansonsten nicht mehr möglich gewesen.

von allen Schülern an räumlich definierten Stationen im Labor vorgenommen werden konnten. Hier bot sich eine Pipettierübung zu Beginn der Arbeitsphase im Laborraum an sowie das Beladen des Agarose-Gels mit DNA-Proben kurz vor der Mittagspause (Abb. 24).

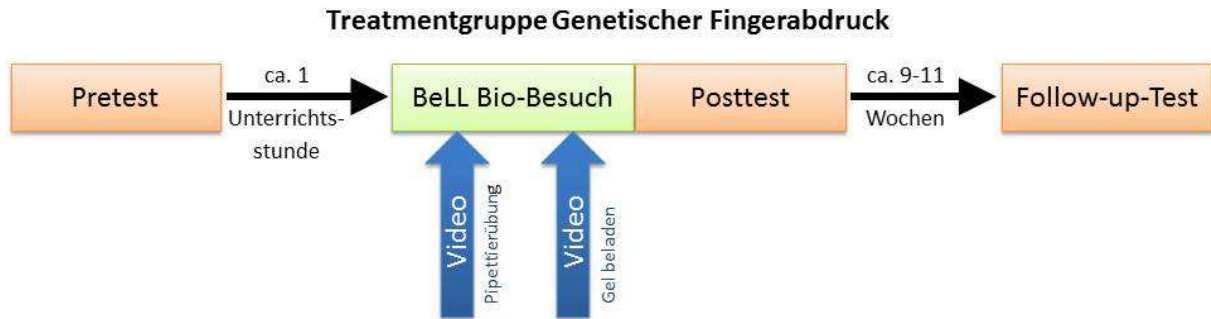


Abbildung 24: Versuchsdesign, der im Rahmen der Hauptstudie durchgeführten Videoanalyse

Um unter Wahrung der Anonymität die Zuordnung der Videodaten zu den im Rahmen der Hauptstudie eingesetzten Fragebögen zu ermöglichen, wurden die Laborkittel mit Nummern am linken Oberarm versehen. Diese waren in den Videoaufnahmen gut zu erkennen, da die Stationen von links gefilmt wurden. Über eine Liste, in der Kittelnummer und Geburtsdatum vermerkt wurden, waren die Videos eindeutig den Fragebögen der Hauptstudie zuzuordnen.

2.2.2.2 Versuchsdesign der Nebenstudie

Das Versuchsdesign der Nebenstudie entspricht weitgehend dem der Hauptstudie. Allerdings wurde hier nur die Treatmentgruppe Genetischer Fingerabdruck untersucht. Als Vergleichsgruppe fungierte die bereits aus der Hauptstudie bekannte Kontrollgruppe. Die Nebenstudie wurde durchgeführt, um ergänzende Informationen über den Einfluss des Schülerlaborbesuchs auf die experimentbezogene Selbstwirksamkeit zu erhalten. Dies gelang durch eine Abwandlung der in der Hauptstudie eingesetzten Messinstrumente (Abb. 25).

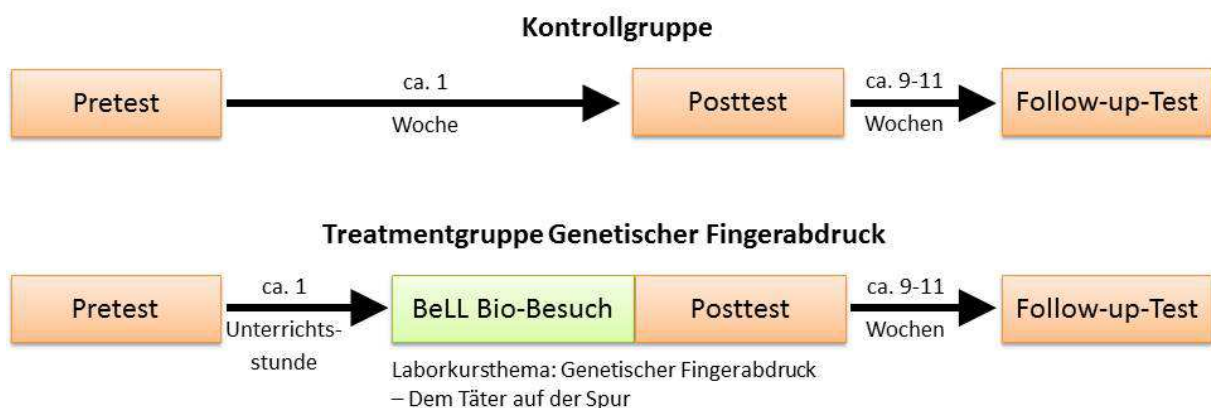


Abbildung 25: Versuchsdesign der Nebenstudie

2.2.2.3 Versuchsdesign der Lehrkräfte-Befragung

Im Gegensatz zur Haupt- und Nebenstudie erfolgte die Fragebogenerhebung bei den Lehrern und Referendaren nicht als wiederholte, sondern als Einmalbefragung. Sie wurde in der Regel unmittelbar im Anschluss an Laborkurse oder Informationsveranstaltungen im BeLL Bio durchgeführt.

2.2.3 Die im Rahmen der Evaluierungen eingesetzten Messinstrumente

Qualitätsmerkmale und Inhalte der in Haupt- und Nebenstudie sowie in der Befragung der Lehrer eingesetzten Messinstrumente werden im Folgenden beschrieben.

2.2.3.1 Qualitätsmerkmale der Messinstrumente

Wird in der Unterrichtsforschung ein Test, eine Einstellungsskala oder ein anderes, auf die Erfassung eines bestimmten Merkmals ausgerichtetes Erhebungsverfahren eingesetzt, so wird dies auch als Messinstrument bezeichnet. In der Unterrichtsforschung werden üblicherweise die Gütekriterien der klassischen Testtheorie herangezogen, um die Qualität eines Messinstrumentes zu belegen (Häußler et al. 1998). Demnach ist ein gutes Messinstrument objektiv, reliabel und valide (Bortz & Döring 2006).

Im Allgemeinen wird ein Test oder Fragebogen dann als objektiv betrachtet, wenn verschiedene Wissenschaftler beim Anwenden des Tests bei denselben Personen zu den gleichen Resultaten gelangen (Bortz & Döring 2006). Die Objektivität wird in der vorliegenden Arbeit durch eine standardisierte, im Abschnitt 2.2.5 dokumentierte Auswerteprozedur gewährleistet (vgl. Bortz & Döring 2006, Bortz & Schuster 2010). Die Verwendung geschlossener Auswahlantworten im Wissenstest sowie der eindeutig abzulesende Grad der Zustimmung in den verwendeten Likert-Skalen sollten dem Kriterium der Objektivität hinreichend genügen (Häußler et al. 1998).

Unter der *Reliabilität* eines Messinstrumentes wird der Grad an Genauigkeit verstanden, mit dem es ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal misst (Lienert & Raatz 1998). Der Reliabilitätskoeffizient gibt den Grad der Reliabilität eines Messinstrumentes an. Üblicherweise – so auch in der vorliegenden Studie – wird Cronbachs α als Reliabilitätskoeffizient berechnet (Zöfel 2002). Wenn ein Messinstrument aus nur wenigen Items besteht, wird eine Reliabilität gelegentlich schon ab einem Reliabilitätskoeffizienten von $\geq 0,5$ als reliabel angesehen (Schön 2007). In der Regel ist aber eine Reliabilität von $\geq 0,7$ erstrebenswert (vgl. Wilde et al. 2009). Die Reliabilitäten der Messinstrumente werden unter Punkt 2.2.3.2.3 (affektiv) und Punkt 2.2.5.3 (kognitiv) vorgestellt. Items, die zu einer deutlichen Verschlechterung der Reliabilität führten, wurden in Orientierung an Häußler et al. (1998) aus der Auswertung entfernt. Dieses Verfahren wird im Weiteren als Reliabilitätsoptimierung bezeichnet.

Die Reliabilität sagt nichts darüber aus, wie gut ein Messinstrument wirklich das misst, was es messen soll. Dies wird erst durch dessen *Validität* gewährleistet (Lienert & Raatz 1998; Bortz & Döring 2006). In der Literatur wird häufig zwischen drei Hauptarten der Validität differenziert: der Inhaltsvalidität, der Kriteriumsvalidität und der Konstruktvalidität. Unter Inhaltsvalidität wird verstanden, dass die Testitems ein zu messendes Konstrukt in seinen wichtigsten Aspekten erfassen. Ein numerisches Maß für die Höhe der Inhaltsvalidität eines Tests gibt es nicht. Sie beruht allein auf subjektiven Einschätzungen (Bortz & Döring 2006). Die Kriteriumsvalidität lässt sich nur bei Vorliegen eines Außenkriteriums prüfen, wenn also bereits eine andere, als gültig anerkannte Beurteilung des untersuchten Persönlichkeitsmerkmals vorliegt (Zöfel 2002). Konstruktvalidität ist gegeben, wenn es gelingt aus dem zu messenden Zielkonstrukt Hypothesen abzuleiten, die von Messwerten bestätigt werden. Zur Bestimmung der Konstruktvalidität wird also nicht ein einziges Außenkriterium herangezogen, sondern ein ganzes Hypothesennetz (Bortz & Döring 2006).

In der vorliegenden Studie kommen viele bewährte Messinstrumente zum Einsatz, die bereits in vergleichbaren Studien erprobt wurden (vgl. Punkt 2.2.3.2.3) und daher als valide angenommen werden. Für diese Studie neu entwickelt wurden die Wissenstests der drei Treatmentgruppen, der Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung (FSWEx) sowie die Kurzskala zur Erfassung der laborbesuchsbedingten Einstellungsänderung zu naturwissenschaftlichen Berufen. Das Messinstrument zur Erfassung der laborbesuchsbedingten Einstellungsänderung gegenüber naturwissenschaftlicher Berufe erhebt nicht den Anspruch das Konstrukt des Berufsinteresses vollständig abzubilden, da dieser Punkt in der vorliegenden Studie lediglich einen Nebenaspekt darstellt. Inhaltsvalidität ist sicherlich insofern gegeben, solange Aussagen aufgrund der mit diesem Messinstrument erhobenen Daten sich nicht auf ein umfassendes Konstrukt, sondern lediglich auf die konkret in den Items formulierten Aussagen beziehen. Zur Einschätzung der Validität der Wissenstests wurde die Korrelation mit dem Außenkriterium der Biologienote in Erwägung gezogen.⁴ Dies erscheint im Falle des Wissenstests der Treatmentgruppe Genetischer Fingerabdruck auch sinnvoll, da die Inhalte des Laborkurses unmittelbar die theoretischen Inhalte des Biologieunterrichtes aufgreifen. Schüler, welche im Biologieunterricht eine höhere kognitive Leistungsfähigkeit zeigen, sollten – Kriteriumsvalidität angenommen – auch im Wissenspretest eine hohe Punktezahl erreichen. Dies wird durch die höchst signifikante Korrelation der Biologienote mit den Testpunkten bestätigt ($r(199)=-0,328$; $p\leq 0,001$). Im Falle des Kurses LDH und noch deutlicher im Fall des Kurses Einzeller erscheint das Außenkriterium der Biologienote nur bedingt geeignet, da die Laborkursinhalte und damit die Inhalte des entsprechenden Wissenstests deutlich über die im Unterricht behandelten Inhalte hinausgehen. Es ist somit nicht davon auszugehen, dass kognitiv leistungsfähigere Schüler im Pretest deutlich besser abschneiden als die

⁴ Auf eine Diskussion der Subjektivität von Schulnoten wird hier verzichtet.

weniger leistungsfähigen. Entsprechend fällt die Korrelation der Testpunkte zur Biologienote in der Treatmentgruppe LDH auch geringer ($r(269)=-0,100$; $p=0,101$) aus. Im Falle des Kurses Einzeller ist keine Korrelation festzustellen ($r(106)=0,004$, $p=0,968$). Die durchgeführten Wissenstests erheben allerdings nicht den Anspruch, Aussagen über die generelle kognitive Leistungsfähigkeit zu ermöglichen. Sie wollen lediglich Aussagen über die aktuelle, themenspezifisch vorhandene kognitive Leistungsfähigkeit treffen. Unter dieser Einschränkung sind die Tests sicherlich als hinreichend valide anzusehen. Zur Begründung der Validität des FSWEx sei hier auf die Diskussion verwiesen.

2.2.3.2 Was wurde in der Hauptstudie erhoben?

Einen Überblick über die in der Hauptstudie eingesetzten Messinstrumente liefert Tabelle 26.

Tabelle 26: Übersicht über die in der Hauptstudie eingesetzten Messinstrumente

Messzeitpunkt	Messinstrumente
Pretest	<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Geburtsdatum, Namenskürzel des Vornamens der Mutter) • Biologienote • Wissenstest (Inhalt abhängig von der Treatmentgruppe) • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeit in der Version FSWEx.V1 (Subskalen: Planung, Durchführung, Auswertung) • Allgemeines biologisches Fähigkeitsselbstkonzept • Fachinteresse (am Unterrichtsfach Biologie) • Sachinteresse (am Experimentieren) • Verständnis (Beitrag von Experimenten zum Verständnis) • Wissenschaftsverständnis
Posttest	<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Geburtsdatum, Namenskürzel des Vornamens der Mutter) • Wissenstest (Inhalt identisch mit Wissenstest in Pre- und Follow-up-Test) • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeit in der Version postFSWEx.V2 (Subskalen: Planung, Durchführung, Auswertung) • Allgemeines biologisches Fähigkeitsselbstkonzept • Alltagsbezug • Authentizität • Betreuung • Offenheit • Zusammenarbeit • Druck • Laboratmosphäre • Aktuelles Interesse (Subskalen: Emotionale Komponente, Wertebezogene Komponente, Epistemische Komponente)
Follow-up-Test	<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Geburtsdatum, Namenskürzel des Vornamens der Mutter) • Wissenstest (Inhalt identisch mit Wissenstest in Pre- und Post-Test) • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeit in der Version FSWEx.V1 (Subskalen: Planung, Durchführung, Auswertung)

Laborvariablen

Messzeitpunkt	Messinstrumente
Follow-up-Test	<ul style="list-style-type: none"> • Allgemeines biologisches Fähigkeitsselbstkonzept • Aktuelles Interesse (Subskalen: Emotionale Komponente, Wertebezogene Komponente, Epistemische Komponente) • Berufsinteresse (Kurzskala zur Erfassung der laborbesuchsbedingten Einstellungsänderung)

2.2.3.2.1 Personenbezogene Daten

Um die Anonymität der befragten Schüler sicherzustellen, wurden in Pre-, Post- und Follow-up-Test lediglich das Geschlecht und das Geburtsdatum erhoben. Um die eindeutige Zuordnung von Pre-, Post- und Follow-up-Test einer Person zu gewährleisten, wurden die ersten beiden Buchstaben des Vornamens der Mutter als weiteres Unterscheidungskriterium erfragt.

2.2.3.2.2 Kognitive Daten (Wissenstest)

Zur Bestimmung des Vorwissens, des Lernzuwachses sowie der langfristigen Behaltensleistung wurde zu allen drei Messzeitpunkten ein Wissenstest im Multiple-Choice-Format eingesetzt. Dieser erfragte in der Treatmentgruppe Genetischer Fingerabdruck allgemeine Grundlagen der modernen Genetik sowie Kenntnisse zu den im Laborkurs angewandten Methoden der DNA-Extraktion, PCR und Gelelektrophorese. Im Wissenstest der Treatmentgruppe LDH wurden Grundlagen der Enzymatik, der Bedeutung der Laktatdehydrogenase im Stoffwechsel sowie der im Labor eingesetzten Methode der Photometrie abgefragt. Der Wissenstest der Treatmentgruppe Einzeller behandelte Grundlagen der Zellbiologie eukaryotischer Einzeller sowie spezielle Fragen zu den im Laborkurs eingesetzten Einzellern *Tetrahymena thermophila* und *Euglena gracilis*. Eine Übersicht über die Gesamtzahl der Aufgaben und Items jedes Tests liefert Tabelle 27.

Tabelle 27: Aufgaben- und Itemzahl der Wissenstests der drei Treatmentgruppen

Treatmentgruppe	Anzahl der Aufgaben	Gesamtzahl der Items
Genetischer Fingerabdruck	23	106
LDH	19	102
Einzeller	21	109

Jede Aufgabe bestand aus vier bis acht Items mit jeweils einer bis mehreren richtigen Antwortmöglichkeiten. Die Aufgaben unterschieden sich zu den drei Testzeitpunkten nicht, sie wurden lediglich in der Reihenfolge variiert. Der Test wurde als *power*- und nicht als *speed*-Test ausgeführt. Während der *speed*-Test Unterschiede in der Bearbeitungsqualität misst die durch Zeitdruck entstehen, erfasst der *power*-Test bei ausreichender Bearbeitungszeit ausschließlich, ob die Aufgaben richtig oder falsch beantwortet wurden (Rost 2004).

Um ausschließen zu können, dass die Bearbeitung der Wissens-Pretests alleine (ohne Laborbesuch) bereits zu einem Wissenszuwachs führt, wurden die Tests ebenfalls in der Kontrollgruppe eingesetzt. Tabelle 28 gibt an, wie viele Schüler der Kontrollgruppe welchen Wissenstest bearbeiteten.

Tabelle 28: Anzahl der Schüler der Kontrollgruppe, die den Wissenstest Genetischer Fingerabdruck, LDH oder Einzeller behandelt haben

Wissenstest	Anzahl der Schüler der Kontrollgruppe
Genetischer Fingerabdruck	31
LDH	70
Einzeller	31

Die unterschiedliche Anzahl der Kontrollgruppenschüler für jeden Wissenstest ist dadurch bedingt, dass für die Bearbeitung des entsprechenden Wissenstests nur Schüler der Kontrollgruppe in Frage kamen, bei denen das Laborkursthema im Untersuchungszeitraum nicht behandelt wurde.

2.2.3.2.3 Affektive Messinstrumente

In der Hauptstudie kamen verschiedene, bereits in anderen Studien bewährte Messinstrumente zum Einsatz. In vielen Fällen musste die Itemzahl der Originalskalen allerdings reduziert werden, da der gesamte Fragebogen – bedingt durch den umfangreichen Wissenstest – nicht zu lang ausfallen sollte. Gemeinsam ist den eingesetzten affektiven Messinstrumenten, dass der Grad der Zustimmung auf einer fünfstufigen Likert-Skala von „stimmt gar nicht“ (Codierung: 0) bis „stimmt völlig“ (Codierung: 4) erhoben wurde. Alle affektiven Messinstrumente werden im Anschluss vorgestellt. Die dort verwendeten Variablennamen entsprechen der Itembezeichnung in der Datenmatrix. Hier steht „pra“ für Pretest affektiv, „poa“ für Posttest affektiv und „foa“ für Follow-up-Test affektiv. Für jedes Messinstrument werden Messzeitpunkte und Reliabilitäten vermerkt. Im Rahmen der Reliabilitätsoptimierung (vgl. Punkt 2.2.3.1) ausgeschlossene Items werden in den folgenden Tabellen grau hinterlegt. Die angegebenen Werte für Cronbachs α beziehen sich auf die Reliabilitäten ohne Berücksichtigung der ausgeschlossenen Items. Differenziert wird – in Anlehnung an die Darstellung unter Punkt 1.2 – in Messinstrumente zur Erfassung von Labor- und Personenvariablen.

Einige der verwendeten Items sind negativ formuliert (z.B. „Die Durchführung der Experimente war langweilig.“). Um sie dennoch gemeinsam mit den übrigen zugehörigen Items ihrer Skalen auswerten zu können, mussten diese umcodiert werden (Bühl 2010). Im Rahmen der Beschreibung der Messinstrumente werden alle umcodierten Items in den Tabellen mit einem Sternchen (*) gekennzeichnet (Tab. 29).

Tabelle 29: Umcodierung der negativ formulierten Items. Diese sind im Weiteren mit „*“ versehen.

Codierung vor dem Umkodieren		Codierung nach dem Umkodieren
0 (stimmt gar nicht)	→	4 (stimmt völlig)
1 (stimmt wenig)	→	3 (stimmt ziemlich)
2 (stimmt teils teils)	→	2 (stimmt teils teils)
3 (stimmt ziemlich)	→	1 (stimmt wenig)
4 (stimmt völlig)	→	0 (stimmt gar nicht)

2.2.3.2.3.1 Messinstrumente zur Erfassung der Personenvariablen

Fachinteresse

Die Skala zur Erfassung des Fachinteresses wurde in Anlehnung an Engeln (2004) erstellt. Sie setzte ihr ebenfalls nur aus zwei Items bestehendes Messinstrument zur Erfassung des Interesses von Schülern am Physikunterricht ein. Das von ihr verwendete Item „Im Physikunterricht fühle ich mich wohl.“ wurde in dieser Untersuchung durch das neue Item „Biologieunterricht finde ich interessant.“ ersetzt, da es direkter das Interesse abfragt (Tab. 30). In der vorliegenden Studie dient das Messinstrument zur Erhebung des Fachinteresses am Biologieunterricht.

Tabelle 30: Items und Reliabilität (Cronbachs α) der Messdimension Fachinteresse im Pretest (Pre) der Hauptstudie

Variablen- name	Item	Pre α
pra4	Biologieunterricht finde ich interessant.	0,893
pra20	Der Biologieunterricht macht mir Spaß.	
Messzeitpunkt: Pretest; Item pra20 in Anlehnung an Engeln (2004), Item pra4 neu		

Sachinteresse

Dieses Messinstrument wird im Pretest eingesetzt, um das Sachinteresse von Schülern am Experimentieren zu erfassen. Während die Items pra6 und pra12 von Engeln (2004) übernommen wurden, wurde Item pra23 selbst konstruiert. Diese Kombination war notwendig, da ein Teil der von Engeln (2004) zur Erhebung des Sachinteresses verwendeten Items nicht das Interesse am Experimentieren, sondern das Interesse an Naturwissenschaften erfasst. Durch Ausschluss des Items pra12 (Tab. 31, grau hinterlegt) konnte die Skalenreliabilität von Cronbachs $\alpha=0,650$ auf Cronbachs $\alpha=0,720$ optimiert werden (vgl. Reliabilitätsoptimierung Punkt 2.2.3.1).

Tabelle 31: Items und Reliabilität (Cronbachs α) der Messdimension Sachinteresse im Pretest (Pre) der Hauptstudie (Das grau hinterlegte Item wurde im Rahmen der Reliabilitätsoptimierung aus der weiteren Auswertung ausgeschlossen und wird im angegebenen Wert für Cronbachs α nicht berücksichtigt.)

Variablen- name	Item	Pre α
pra6	Experimente durchzuführen, macht mir einfach Spaß.	0,720
pra12	Wenn ich Experimente durchführen kann, bin ich bereit auch Freizeit dafür zu verwenden.	
pra23	Experimente durchzuführen finde ich wichtig.	
Messzeitpunkt: Pretest; Item pra6 & 12 in Anlehnung an Engeln (2004); Item pra23 neu		

Verständnis

Mit dem Messinstrument Verständnis wird erhoben, wie Schüler den Beitrag von Experimenten zum Verständnis naturwissenschaftlicher Theorien und Fragestellungen einschätzen. Item pra28 wurde in Anlehnung an Pawek (2009) formuliert. Dieser verwendet das Item allerdings als Bestandteil seiner Skala zur Erfassung des akademischen Fähigkeitsselbstkonzeptes. Item pra7 und pra19 wurden neu formuliert (Tab. 32).

Tabelle 32: Items und Reliabilität (Cronbachs α) der Messdimension Verständnis im Pretest (Pre) der Hauptstudie

Variablen- name	Item	Pre α
pra7	Experimente helfen mir, biologische Fragestellungen zu verstehen.	0,714
pra19	Die Durchführung von Experimenten trägt zum Verständnis bei.	
pra28	Anhand anschaulicher Experimente verstehe ich auch komplizierte Theorien.	
Messzeitpunkt: Pretest; Item pra28 in Anlehnung an Pawek (2009), Item pra7 & 19 neu		

Wissenschaftsverständnis

Die Items des Messinstruments „Wissenschaftsverständnis“ entstammen einem von Urhahne und Hopf (2004) ins Deutsche übersetzten Fragebogen zur Erfassung epistemologischer Überzeugungen. Das englischsprachige Original wurde von Conley et al. (2004) entwickelt. In vier Subskalen „Quelle des Wissens“, „Sicherheit des Wissens“, „Entwicklung des Wissens“ und „Rechtfertigung des Wissens“ erhebt dieser Fragebogen die Vorstellungen von Lernenden darüber, wie wissenschaftliche Erkenntnisse gewonnen werden und mit welcher Sicherheit diese Erkenntnisse bestehen. Die vier hier verwendeten Items (Tab. 33) sind Bestandteil der Subskala „Rechtfertigung des Wissens“. Sie erfassen die Einstellung zur Bedeutung von Experimenten im Erkenntnisprozess. Im deutschsprachigen Original besteht die Subskala aus neun Items und weist eine Reliabilität von

Cronbachs $\alpha=0,68$ auf (Urhahne & Hopf 2004). Die Verkürzung der Skala auf vier Items bedingte eine Verringerung der Reliabilität auf Cronbachs $\alpha=0,55$.

Tabelle 33: Items und Reliabilität (Cronbachs α) der Messdimension Wissenschaftsverständnis im Pretest (Pre) der Hauptstudie

Variablen-name	Item	Pre α
pra10	Es ist wichtig, Experimente mehr als einmal durchzuführen, um Ergebnisse abzusichern.	0,547
pra25	Die Ideen zu naturwissenschaftlichen Experimenten kommen daher, dass man neugierig ist und darüber nachdenkt, wie etwas funktioniert.	
pra30	Gute Theorien stützen sich auf die Ergebnisse aus vielen verschiedenen Experimenten.	
pra32	In den Naturwissenschaften kann es mehrere Wege geben, um Vorstellungen zu überprüfen.	
Messzeitpunkt: Pretest; Items aus Urhahne & Hopf (2004)		

Allgemeines Biologisches Fähigkeitsselbstkonzept

Die Items zur Erhebung des biologischen Fähigkeitsselbstkonzeptes wurden aus der Arbeit von Engeln (2004) übernommen und geringfügig adaptiert. Engeln (2004) nutzte die Items, um das naturwissenschaftliche Selbstkonzept zu messen. In der vorliegenden Untersuchung finden die Items dagegen Verwendung, um konkreter das biologische Selbstkonzept zu erfassen. Die Items wurden in identischer Form in Pre-, Post- und Follow-up-Test von Treatmentgruppen und Kontrollgruppe eingesetzt (Tab. 34).

Tabelle 34: Items und Reliabilitäten (Cronbachs α) der Messdimension Allgemeines biologisches Fähigkeitsselbstkonzept in den Schülerfragebögen der Pre-(Pre), Post- (Post) und Follow-up-Tests (Fol), *Items wurden umkodiert

Variablen- name(n)	Item	Pre α	Post α	Fol α
pra1*, poa26*, foa23*	Bei manchen Sachen in der Biologie, die ich nicht verstanden habe, weiß ich von vornherein: „Das verstehe ich nie.“	0,839	0,839	0,858
pra14*, poa41*, foa28*	Für Biologie habe ich einfach keine Begabung.			
pra21*, poa45*, foa11*	Biologie liegt mir nicht besonders.			
Messzeitpunkt: Pre-, Post- und Follow-up-Test; Items in Anlehnung an Engeln (2004)				

Aktuelles Interesse

Die im Pretest eingesetzten Messinstrumente dienten dazu, das bereits vor dem Laborbesuch bestehende Fach- und Sachinteresse der Schüler zu erfassen. Um auch das aktuelle, also das während des Laborbesuches bestehende Interesse messen zu können, wurden im Posttest die drei Subskalen *Emotionales Interesse* (Tab. 35), *Wertebezogenes Interesse* (Tab. 36) und *Epistemisches Interesse* (Tab. 37) verwendet. Diese wurden weitgehend der Arbeit von Engeln (2004) entnommen und modifiziert. Während die zuerst genannte Subskala die emotionale Aktivierung der Schüler während des Experimentierens erfragt, bezieht sich die Subskala *Wertebezogenes Interesse* auf die den Experimenten persönlich beigemessene Relevanz. In der Subskala *Epistemisches Interesse* wird die Bereitschaft erfragt, sich noch über den Laborbesuch hinausgehend mit dessen Inhalten zu beschäftigen. Im Follow-up-Test wurden die Items der ersten beiden Subskalen in fast identischer Form verwendet. Hier wurde ergänzend lediglich noch einmal ausdrücklich darauf hingewiesen, dass sich die Items auf den BeLL Bio-Besuch beziehen. In Anlehnung an Engeln (2004) wurden die Follow-up-Test Items der Subskala *Epistemisches Interesse* so adaptiert, dass die Schüler rückblickend beurteilen mussten, ob sie ihre Absichten, sich nach dem BeLL Bio-Besuch noch mit dessen Inhalten zu beschäftigen, tatsächlich umgesetzt haben (Tab. 37). Für die Kontrollgruppe mussten die Items entsprechend so angepasst werden, dass sie sich nicht auf den Laborbesuch, sondern auf den schulischen Biologieunterricht beziehen. Eine Übersicht dieser Items findet sich im Anhang A.

Tabelle 35: Items und Reliabilitäten (Cronbachs α) der Messdimension der emotionalen Komponente des aktuellen Interesses in Post-(Post) und Follow-up-Test (Fol) der Hauptstudie, *Items wurden umkodiert

Variablen-name	Item	Post α	Fol α
poa6	Die Experimente haben Spaß gemacht.	0,854	0,889
foa14	Die Experimente am Tag des BeLL Bio-Besuches haben Spaß gemacht.		
poa24*	Die Durchführung der Experimente war langweilig.		
foa20*	Die Durchführung der Experimente im BeLL Bio war langweilig.		
poa34	Ich fand die Experimente spannend.		
foa30	Ich fand die Experimente im BeLL Bio spannend.		
Messzeitpunkt: Post- & Follow-up-Test; Items in Anlehnung an Engeln (2004)			

Tabelle 36: Items und Reliabilitäten (Cronbachs α) der Messdimension der wertebezogenen Komponente des aktuellen Interesses in Post-(Post) und Follow-up-Test (Fol) der Hauptstudie

Variablen-name	Item	Post α	Fol α
poa8	Die heute durchgeführten Experimente erscheinen mir wichtig.	0,746	0,776
foa4	Die im BeLL Bio durchgeführten Experimente erscheinen mir wichtig.		
poa30	Dass wir heute Experimente durchgeführt haben, erscheint mir sinnvoll.		
foa26	Dass wir im BeLL Bio Experimente durchgeführt haben, erscheint mir sinnvoll.		
poa36	Dass wir heute Experimente durchgeführt haben, ist mir persönlich wichtig.		
foa12	Dass wir im BeLL Bio Experimente durchgeführt haben, ist mir persönlich wichtig.		
Messzeitpunkt: Post- & Follow-up-Test; Items in Anlehnung an Engeln (2004)			

Tabelle 37: Items und Reliabilitäten (Cronbachs α) der Messdimension der epistemischen Komponente des aktuellen Interesses in Post-(Post) und Follow-up-Test (Fol) der Hauptstudie

Variablen-name	Item	Post α	Fol α
poa28	Ich werde in Büchern/im Internet nachlesen, um mehr Informationen über das im BeLL Bio behandelte Thema zu erhalten.	0,728	0,720
foa31	Ich habe nach dem BeLL Bio-Besuch in Büchern/im Internet nachgelesen, um mehr Informationen über das im BeLL Bio behandelte Thema zu erhalten.		
poa33	Ich werde außerhalb des Unterrichts über Dinge nachdenken, die wir im BeLL Bio gesehen oder angesprochen haben.		
foa7	Ich habe nach dem BeLL Bio-Besuch außerhalb des Unterrichts über Dinge nachgedacht, die wir im BeLL Bio gesehen oder angesprochen haben.		
poa42	Ich werde mit Freunden, Eltern oder Geschwistern über Dinge sprechen, die ich im BeLL Bio erlebt habe.		
foa18	Ich habe nach dem BeLL Bio-Besuch mit Freunden, Eltern oder Geschwistern über Dinge gesprochen, die ich im BeLL Bio erlebt habe.		
Messzeitpunkt: Post- & Follow-up-Test; Items in Anlehnung an Engeln (2004)			

Fragebogen zur subjektiv wahrgenommenen Experimentierkompetenz

Ein im Rahmen dieser Untersuchung neu entstandenes Messinstrument ist der Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung (FSWEx). Der FSWEx wurde auf der Grundlage des Modells experimenteller Kompetenz nach Schreiber et al. (2009) entwickelt. Er erfasst in den drei Subskalen Planung, Durchführung und Auswertung die von Schülern in Bezug auf ihre eigene Person eingeschätzte Kompetenz, Experimente zu planen, durchzuführen und auszuwerten. Ausführlich wird der Entwicklungsprozess unter Punkt 3.2.1 dieser Arbeit vorgestellt. Der FSWEx existiert in den drei Versionen FSWEx.V1, FSWEx.V2 und postFSWEx.V1. Die Kurzversion FSWEx.V1 kam in Pre- und Follow-up-Test der Hauptstudie zum Einsatz.⁵ Der postFSWEx.V1 fand im Posttest der Hauptstudie Verwendung und der FSWEx.V2 wurde zu allen drei Messzeitpunkten der Nebenstudie eingesetzt. Der FSWEx.V1 erfasst die experimentbezogene Selbstwirksamkeitserwartung situationsunabhängig. Personen, die das Messinstrument ausfüllen, müssen also zuvor keine konkrete Experimentiersituation durchlaufen haben. Die Tabellen 38 bis 40 geben einen Überblick über die Items der drei Subskalen des FSWEx.V1.

Tabelle 38: Items und Reliabilitäten (Cronbachs α) der FSWEx.V1-Subskala Planung in Pre-(Pre) und Follow-up-Test (Fol) der Hauptstudie

Variablen- name	Item	Pre α	Fol α
pra11, foa5	Durch Beobachtungen in meinem täglichen Leben fallen mir häufig Fragen ein, die man durch ein Experiment aufklären könnte.	0,513	0,629
pra16, foa13	Es gelingt mir, anhand theoretischer Überlegungen, Erwartungen (Hypothesen) über den Ausgang eines Experimentes zu formulieren.		
pra27, foa6	Es fällt mir leicht, ein Experiment zur Aufklärung einer vorgegebenen Fragestellung zu entwickeln.		
Messzeitpunkt: Pre- & Follow-up-Test; Items neu			

⁵ Im Rahmen der Hauptstudie wurde der FSWEx.V1 entwickelt und zur Reliabilitätsoptimierung um weitere Items ergänzt. Resultierend daraus kam bei den Schülern der Kontrollgruppe bereits der erweiterte FSWEx.V2 zu Einsatz. Um eine gemeinsame Auswertung zu ermöglichen, fließen in die Auswertung von Post- und Follow-up-Test der Hauptstudie für alle Schüler nur die Items des FSWEx.V1 ein. In der Nebenstudie wurde zu allen drei Messzeitpunkten der FSWEx.V2 eingesetzt und ausgewertet.

Tabelle 39: Items und Reliabilitäten (Cronbachs α) der FSWEx.V1-Subskala Durchführung in Pre-(Pre) und Follow-up-Test (Fol) der Hauptstudie (Das grau hinterlegte Item wurde im Rahmen der Reliabilitätsoptimierung aus der Auswertung ausgeschlossen und wird im angegebenen Wert für Cronbachs α nicht berücksichtigt, *Items wurden umkodiert)

Variablen- name	Item	Pre α	Fol α
pra31*, foa22*	Für die Durchführung von Experimenten habe ich kein Händchen.	0,576	0,629
pra8, foa27	Im Umgang mit Laborgeräten bin ich gut.		
pra22*, foa32*	Mit dem Aufschreiben von Versuchsbeobachtungen habe ich Schwierigkeiten.		
Messzeitpunkt: Pre- & Follow-up-Test; Item pra31/foa22 aus Pawek (2009), Items pra8/foa27 & pra22/foa32 neu			

Tabelle 40: Items und Reliabilitäten (Cronbachs α) der FSWEx.V1-Subskala Auswertung in Pre-(Pre) und Follow-up-Test (Fol) der Hauptstudie, *Items wurden umkodiert

Variablen-name	Item	Pre α	Fol α
pra13, foa9	Die Auswertung von Messdaten fällt mir leicht.	0,647	0,762
pra26*, foa24*	Ich habe häufig Schwierigkeiten bei der Interpretation von Versuchsergebnissen.		
pra18, foa17	Es ist kein Problem für mich, Versuchsbeobachtungen zu deuten.		
Messzeitpunkt: Pre- & Follow-up-Test; Items neu			

Im Gegensatz zum FSWEx.V1 erfasst der postFSWEx.V1 die experimentbezogene Selbstwirksamkeit situationsabhängig. Der postFSWEx.V1 wurde in der Hauptstudie im Posttest eingesetzt, um die experimentbezogene Selbstwirksamkeit, bezogen auf die unmittelbar vorausgehende Experimentiersituation im BeLL Bio, einzuschätzen (Tab. 41-43).

Tabelle 41: Items und Reliabilität (Cronbachs α) der postFSWEx.V1-Subskala Planung im Posttest der Hauptstudie

Variablen- name	Item	Post α
poa46	Die durchgeführten Experimente haben bei mir Fragen aufgeworfen, die ich mit anderen Experimenten beantworten könnte.	0,620
poa38	Ich hatte mir schon vor dem Experiment überlegt, was herauskommen müsste und lag mit meiner Überlegung (Hypothese) richtig.	
poa29	Die Planung der durchgeführten Experimente wäre mir leicht gefallen.	
Messzeitpunkt: Posttest; Items neu		

Tabelle 42: Items und Reliabilität (Cronbachs α) der postFSWEx.V1-Subskala Durchführung im Posttest der Hauptstudie (Das grau hinterlegte Item wurde im Rahmen der Reliabilitätsoptimierung aus der Auswertung ausgeschlossen und wird im angegebenen Wert für Cronbachs α nicht berücksichtigt, *Items wurden umkodiert)

Variablen- name	Item	Post α
poa25	Die Durchführung der Experimente ist mir leicht gefallen.	0,582
poa40	Der Umgang mit den Laborgeräten war kein Problem.	
poa32*	Mit dem Aufschreiben der Versuchsbeobachtungen hatte ich Schwierigkeiten.	
Messzeitpunkt: Posttest; Items neu		

Tabelle 43: Items und Reliabilität (Cronbachs α) der postFSWEx.V1-Subskala Auswertung im Posttest der Hauptstudie, *Items wurden umkodiert

Variablen- name	Item	Post α
poa35	Die Auswertung der Versuchsergebnisse war leicht.	0,759
poa17*	Ich hatte Schwierigkeiten bei der Interpretation der Versuchsergebnisse.	
poa43	Es war kein Problem für mich, die Versuchsbeobachtungen zu deuten.	
Messzeitpunkt: Posttest; Items neu		

Messinstrumente zur Erfassung von Personenmerkmalen

In Ergänzung zu den bereits erwähnten Messinstrumenten wurden im Follow-up-Test weiterhin der Migrationshintergrund, die beruflichen Pläne nach Abschluss der Schule sowie die durch den BeLL Bio-Besuch erzielte Einstellungsänderung gegenüber naturwissenschaftlichen Berufen erfragt. Der Migrationshintergrund wurde durch die in Abbildung 26 angeführten Fragen erhoben. Als Schüler mit Migrationshintergrund wurden diejenigen eingestuft, die mindestens in einer der beiden Fragen „zugewandert“ ankreuzten.

Sind Sie in Deutschland geboren oder sind Sie aus einem anderen Land zugewandert?	
<input type="checkbox"/> in Deutschland geboren	<input type="checkbox"/> zugewandert
Sind Ihre Eltern in Deutschland geboren?	
<input type="checkbox"/> in Deutschland geboren	<input type="checkbox"/> zugewandert

Abbildung 26: Frage zur Erfassung des Migrationshintergrundes

Die berufliche Orientierung wurde in folgender Form erhoben (Abb. 27):

Wie sehen Ihre Pläne nach Abschluss der Schule aus?

☐ Ich strebe einen Ausbildungsplatz im naturwissenschaftlichen Bereich an.
Wenn ja: In welcher naturwissenschaftlichen Fachrichtung?

☐ Biologie
 ☐ Chemie
 ☐ Physik

☐ Technik
 ☐ Sonstige _____

☐ Ich strebe einen Ausbildungsplatz in einem nicht naturwissenschaftlichen Bereich an.

☐ Ich strebe ein Studium im naturwissenschaftlichen Bereich an.
Wenn ja: In welcher naturwissenschaftlichen Fachrichtung?

☐ Biologie
 ☐ Chemie
 ☐ Physik

☐ Technik
 ☐ Sonstige _____

☐ Ich strebe ein Studium in einem nicht naturwissenschaftlichen Bereich an.

☐ Ich habe andere Pläne.

Abbildung 27: Frage zur Erfassung der beruflichen Orientierung

Diejenigen Schüler, die angaben, einen naturwissenschaftlichen Ausbildungsplatz oder ein naturwissenschaftliches Studium absolvieren zu wollen, werden im weiteren Verlauf als Personen mit naturwissenschaftlicher Berufsorientierung bezeichnet. Die übrigen Schüler werden als nicht naturwissenschaftlich orientiert eingestuft.

Tabelle 44 gibt einen Überblick über die im Follow-up-Test eingesetzten Items zur Erfassung der durch den Laborbesuch bedingten Einstellungsänderung zu naturwissenschaftlichen Berufen. Die Items wurden in dieser Studie neu entwickelt.

Tabelle 44: Items und Reliabilität (Cronbachs α) der Messdimension Berufsinteresse im Follow-up-Test(Fol) der Hauptstudie

Variablen- name	Item	Fol α
foa1	Nach dem BeLL Bio-Besuch kann ich mir eher vorstellen, eine naturwissenschaftliche Ausbildung/ein naturwissenschaftliches Studium anzutreten, als vor dem Besuch.	0,874
foa15	Der BeLL Bio-Besuch hat mein Interesse für naturwissenschaftliche Berufe geweckt.	
foa19	Der BeLL Bio-Besuch hat mich angeregt, mich über naturwissenschaftliche Berufe zu informieren.	
Messzeitpunkt: Follow-up-Test; Items neu		

2.2.3.2.3.2 Messinstrumente zur Erfassung der Laborvariablen

Die im Posttest erhobenen Laborvariablen erfassen die Wahrnehmung der Lernumgebung Labor durch die Schüler der Treatmentgruppen. In der Kontrollgruppe wurden die Items geringfügig adaptiert, so dass nicht die Wahrnehmung der Lernumgebung Labor, sondern die der Lernumgebung schulischer Biologieunterricht erfasst wurde. Dies geschah in der Regel durch Einfügen der Worte „im Biologieunterricht“. Aus dem Item „Ich konnte eine Bedeutung der durchgeführten Experimente für das alltägliche Leben erkennen“ wurde beispielsweise das Item „Ich kann eine Bedeutung der im Biologieunterricht durchgeführten Experimente für das alltägliche Leben erkennen“. Eine ausführliche Auflistung der in der Kontrollgruppe adaptierten Items befindet sich im Anhang A. Im Folgenden wird die Itemversion der Treatmentgruppen vorgestellt.

Alltagsbezug

Der von den Schülern empfundene Alltagsbezug der Lernumgebung BeLL Bio wird durch drei Items erhoben, die der Studie von Pawek (2009) entnommen wurden. Das ursprüngliche Messinstrument besteht aus vier Items und bezieht sich auf Naturwissenschaften allgemein. In den Items poa5 und poa15 wurde in der vorliegenden Studie das Wort „Naturwissenschaften“ durch „Biologie“ ersetzt (Tab. 45).

Tabelle 45: Items und Reliabilität (Cronbachs α) der Messdimension Alltagsbezug im Posttest (Post) der Hauptstudie

Variablen- name	Item	Post α
poa2	Ich konnte eine Bedeutung der durchgeführten Experimente für das alltägliche Leben erkennen.	0,778
poa5	Ich habe heute etwas über die Bedeutung der Biologie für unseren Alltag gelernt.	
poa15	Ich habe heute etwas über die Bedeutung der Biologie für unsere Gesellschaft gelernt.	
Messzeitpunkt: Posttest; Items in Anlehnung an Pawek (2009)		

Authentizität

Wie authentisch Schüler die Laborkurse im BeLL Bio empfinden, erheben die in Tabelle 46 aufgelisteten Items. Item poa10 und poa21 stammen aus einem Messinstrument zur Erfassung der Authentizität von Engeln (2004). Item poa3 wurde neu formuliert.

Tabelle 46 Items und Reliabilität (Cronbachs α) der Messdimension Authentizität im Posttest (Post) der Hauptstudie

Variablen- name	Item	Post α
poa3	Ich habe heute etwas darüber gelernt, wie Wissenschaftler arbeiten.	0,807
poa10	Ich habe heute ein Gefühl dafür bekommen, wie Forschung funktioniert.	
poa21	Ich habe heute einen Einblick in den Berufsalltag von Wissenschaftlern bekommen.	
Messzeitpunkt: Posttest; Items poa10 & 21 in Anlehnung an Engeln (2004), Item poa3 neu		

Betreuer

Als Experte stellt der Kursbetreuer in der Lernumgebung Labor eine wichtige Bezugsperson dar. Mittels der drei in Tabelle 47 angeführten Items sollte die dem Betreuer beigemessene Bedeutung erfasst werden. Die Items stammen aus der vier Items umfassenden Skala Betreuung/Atmosphäre von Pawek (2009). Aufgrund der geringen Reliabilität von Cronbachs $\alpha=0,454$ wurde die Skala verworfen. Item poa4 fließt als Einzelitem in die Auswertung ein.

Tabelle 47: Items der Messdimension Betreuer im Posttest (Post) der Hauptstudie (Grau hinterlegte Items wurden – bedingt durch eine geringe Reliabilität der Gesamtskala – aus der weiteren Auswertung ausgeschlossen. Item poa4 fließt als Einzelitem in die Auswertung ein.)

Variablen- name	Item	Post α
poa4	Die Gespräche mit den Betreuern waren mir wichtig.	---
poa11	Ich hatte die Möglichkeit, den Betreuern des BeLL Bio Fragen zu stellen.	
poa19	Ich habe das Gefühl, dass die Betreuer von Naturwissenschaften/Technik fasziniert sind.	
Messzeitpunkt: Posttest; Items in Anlehnung an Pawek (2009)		

Zusammenarbeit

Zur Erfassung der Qualität der Zusammenarbeit der Schüler untereinander wurden die in Tabelle 48 aufgeführten Items verwendet. Diese kamen in ähnlicher Form bereits bei Engeln (2004) und Pawek (2009) zum Einsatz.

Tabelle 48: Items und Reliabilität (Cronbachs α) der Messdimension Zusammenarbeit im Posttest (Post) der Hauptstudie

Variablen- name	Item	Post α
poa9	Ich habe während des Experimentierens mit meinen Mitschülerinnen und Mitschülern über das Experiment diskutiert.	0,519
poa14	Ich habe während der Experimente meinen Mitschülerinnen/Mitschülern etwas erklärt oder mir ist etwas erklärt worden.	
poa22	Ich habe während des Experimentierens gut mit meinen Mitschülerinnen und Mitschülern im Team zusammengearbeitet.	
Messzeitpunkt: Posttest; Items in Anlehnung an Engeln (2004) & Pawek (2009)		

Druck/Anspannung

Hoher Leistungsdruck wird als negativer Prädiktor für intrinsische Motivation gesehen (vgl. Punkt 1.2.2.7). Das in Tabelle 49 aufgeführte Messinstrument erhebt den im Labor empfundenen Druck in zwei Items. Diese wurden in Anlehnung an die Subskala Druck/Anspannung der Kurzsкала zur intrinsischen Motivation konstruiert. Wilde et al. (2009) entwarfen diese Kurzsкала auf Grundlage des englischsprachigen *Intrinsic Motivation Inventory* von Deci & Ryan (2012).

Tabelle 49: Items und Reliabilität (Cronbachs α) der Messdimension Druck/Anspannung im Posttest (Post) der Hauptstudie.

Variablen- name	Item	Post α
poa23	Während der Laborarbeit fühlte ich mich unter Leistungsdruck gesetzt.	0,599
poa39	Ich fühlte mich bei der Laborarbeit angespannt.	
Messzeitpunkt: Posttest; Items in Anlehnung an Wilde et al. (2009), Deci & Ryan (2012)		

Offenheit

Um den Grad der von den Schülern wahrgenommenen Offenheit der Lernumgebung BeLL Bio zu erfassen, wurden die Items der Tabelle 50 verwendet. Die drei Items wurden von Engeln (2004) übernommen. Item poa27 und poa12 wurden geringfügig umformuliert. Die Reliabilität der Gesamtskala von Cronbachs $\alpha=0,521$ konnte durch Ausschluss des Items poa7 auf Cronbachs $\alpha=0,703$ gesteigert werden.

Tabelle 50: Items und Reliabilität (Cronbachs α) der Messdimension Offenheit im Posttest (Post) der Hauptstudie (Das grau hinterlegte Item wurde im Rahmen der Reliabilitätsoptimierung aus der Auswertung ausgeschlossen und wird im angegebenen Wert für Cronbachs α nicht berücksichtigt, *Items wurden umkodiert)

Variablen -name	Item	Post α
poa7*	Während des Experimentierens hatte ich das Gefühl, nichts selber bestimmen zu können.	0,703
poa12	Während des Experimentierens hatte ich die Möglichkeit, eigene Ideen auszuprobieren.	
poa27	Während der Experimente konnte ich eigene Entscheidungen treffen.	
Messzeitpunkt: Posttest; Items in Anlehnung an Engeln (2004)		

Laboratmosphäre

Von Pawek (2009) wurde in der Skala Betreuung/Atmosphäre das Item „Die Arbeitsatmosphäre während des Experimentierens fand ich gut.“ verwendet. In der vorliegenden Studie sollte nicht nur die Atmosphäre während des Experimentierens, sondern die gesamte laborraumbedingte Atmosphäre erfasst werden. Somit wurden die in Tabelle 51 dargestellten Items neu entwickelt. Wie die Reliabilität von Cronbachs $\alpha=0,169$ für die Gesamtskala zeigt, bilden diese kaum ein gemeinsames Konstrukt ab. Da Item poa44 explizit die Atmosphäre im Laborraum erfragt, wird es als Einzelitem in der weiteren Auswertung berücksichtigt.

Tabelle 51: Items der Messdimension Laboratmosphäre im Posttest (Post) der Hauptstudie (Das grau hinterlegte Item wurde – bedingt durch eine geringe Reliabilität der Gesamtskala – aus der weiteren Auswertung ausgeschlossen. Item poa44 fließt als Einzelitem in die Auswertung ein, *Items wurden umkodiert)

Variablen- name	Item	Post α
poa31*	Der Laborraum wirkte beunruhigend auf mich.	---
poa44	Der Laborraum verbreitete eine angenehme Atmosphäre.	
Messzeitpunkt: Posttest; Items neu konstruiert		

Transparenz der Versuchsanleitung

Die folgenden Items in Tabelle 52 wurden eingesetzt, um zu erheben, wie transparent das Versuchsskript von den Schülern wahrgenommen wurde. Da in der Kontrollgruppe keine Versuchsanleitungen ausgegeben wurden, kamen diese Items auch nur in den Fragebögen der Treatmentgruppen zum Einsatz.

Tabelle 52: Items und Reliabilität (Cronbachs α) der Messdimension Transparenz im Posttest (Post) der Hauptstudie

Variablen- name	Item	Post α
poAnl1	Ich wusste an jeder Stelle der Versuchsanleitung, was ich zu tun hatte.	0,803
poAnl2	Die Aufgaben in der Versuchsanleitung waren klar formuliert.	
poAnl3	Ich habe die Anleitungen zum Experimentieren gut verstanden.	
Messzeitpunkt: Posttest; Items neu		

2.2.3.2.4 Videoaufzeichnungen

Als ergänzende Messinstrumente wurden für den Laborkurs Genetischer Fingerabdruck die im Rahmen der BeLL Bio-Videostudie (Dahmen 2012) erzeugten Videos vom Experimentierverhalten der Schüler verwendet. Mit zwei Videokameras wurde das Verhalten der Schüler bei einer Pipettierübung sowie beim Beladen eines Agarose-Gels aufgezeichnet. Angeboten wurden die Experimentiertätigkeiten an vier Stationen. Die Kameras wurden in etwa zweieinhalb Metern Höhe im Laborraum stationär platziert, so dass die Experimentiertätigkeit von jedem Schüler deutlich im Video erkennbar war (Abb. 28).

Die Aufnahmefunktion der Kameras wurde so lange aktiviert, bis alle Kursteilnehmer die Pipettierübung, bzw. das Beladen des Gels absolviert hatten. Die so aufgezeichneten, jeweils etwa 20minütigen, Videosequenzen zeigten nacheinander Schüler bei der entsprechenden Experimentiertätigkeit.

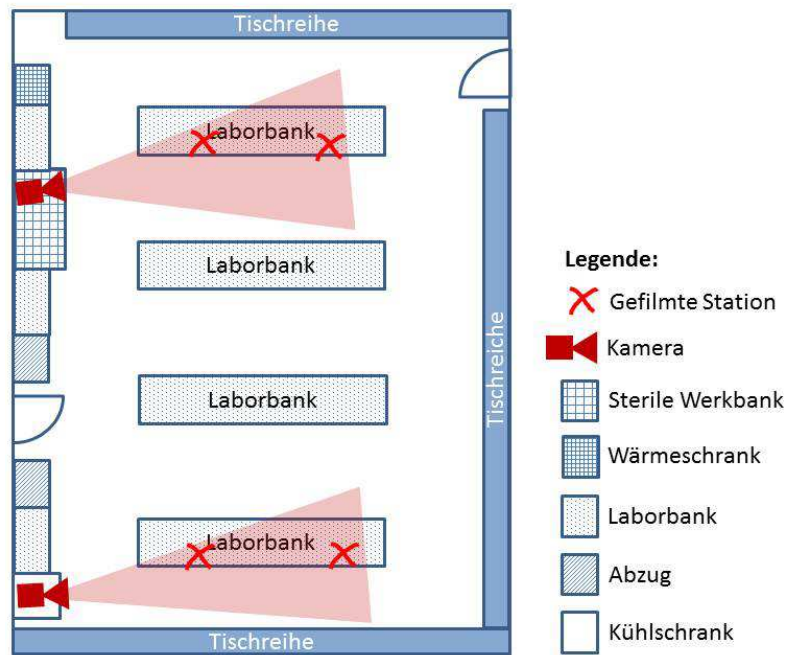


Abbildung 28: Position der beiden Videokameras im Laborraum des BeLL Bio

2.2.3.3 Was wurde in der Nebenstudie erhoben?

In der Hauptstudie wurden verschiedenste affektive und kognitive Messinstrumente eingesetzt. Unter anderem fand dort im Pre- und Follow-up-Test der FSWEx.V1 Verwendung. Im Posttest wurde der situationsabhängige postFSWEx.V1 eingesetzt, um die unmittelbar beim Laborbesuch empfundene experimentbezogene Selbstwirksamkeit zu erfassen. Da sich die Items des postFSWEx.V1 deutlich von denen des FSWEx.V1 unterscheiden, ist eine gemeinsame Auswertung über alle drei Messzeitpunkte allerdings nur korrelativ möglich. Aussagen über eine möglicherweise kurzfristige, durch den Laborbesuch bedingte quantitative Veränderung der experimentbezogenen Selbstwirksamkeit können so nicht getroffen werden. Da eine etwaige Änderung ein enormes Potential des Laborbesuches belegen könnte, wurde in der Nebenstudie ein identischer FSWEx zu allen drei Messzeitpunkten verwendet. Zum Zeitpunkt der Nebenstudie stand bereits der erweiterte FSWEx.V2 zur Verfügung. Zusätzlich wurden die zuvor unter Punkt 2.2.3.2.1 beschriebenen personenbezogenen Daten erhoben. In Tabelle 53 sind die in der Nebenstudie eingesetzten Messinstrumente zusammengestellt.

Tabelle 53: Übersicht der in der Nebenstudie eingesetzten Messinstrumente

Messzeitpunkt	Messinstrumente
Pretest	<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Geburtsdatum, Namenskürzel des Vornamens der Mutter) • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung in den Subskalen Planung, Durchführung und Auswertung (FSWEx.V2)
Posttest	<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Geburtsdatum, Namenskürzel des Vornamens der Mutter) • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung in den Subskalen Planung, Durchführung und Auswertung (FSWEx.V2)
Follow-up-Test	<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Geburtsdatum, Namenskürzel des Vornamens der Mutter) • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung in den Subskalen Planung, Durchführung und Auswertung (FSWEx.V2)

2.2.3.3.1 Fragebogen zur subjektiv wahrgenommenen Experimentierkompetenz

Der FSWEx.V2 ergänzt die Subskalen des FSWEx.V1 um jeweils drei weitere Items. Diese ermöglichen eine umfassendere Abdeckung des Konstrukts der Experimentierkompetenz nach dem Modell Schreibers et al. (2009) und optimierten die Reliabilität des Messinstrumentes. Die Tabellen 54 bis 56 beinhalten die Items, um welche der FSWEx.V2 gegenüber dem FSWEx.V1 (vgl. Tabellen 38 bis 40) ergänzt wurde.

Tabelle 54: Items und Reliabilität (Cronbachs α) der FSWEx.V2-Subskala Planung im Pre-, Post- und Follow-up-Test der Nebenstudie (Angabe sind nur die Items, um die der FSWEx.V2 den FSWEx.V1 ergänzt. Die Reliabilitäten beziehen sich auf die vollständige Subskala.)

Variablen-name	Item	Pre α	Post α	Fol α
pra9, poa20, foa9	Ich bin gut darin, mir zu überlegen, was für Geräte/Materialien ich für ein Experiment benötige.	0,789	0,833	0,849
pra15, poa13, foa10	Es ist einfach für mich, mir eine Experimentieranleitung für die Beantwortung einer bestimmten Fragestellung zu überlegen.			
pra29, poa48, foa25	Wenn ich eine Idee für ein Experiment habe, fällt es mir leicht, mir einen geeigneten Versuchsaufbau dafür auszudenken.			

Tabelle 55: Items und Reliabilität (Cronbachs α) der FSWEx.V2-Subskala Durchführung im Pre-, Post- und Follow-up-Test der Nebenstudie (Angabe sind nur die Items, um die der FSWEx.V2 den FSWEx.V1 ergänzt. Die Reliabilitäten beziehen sich auf die vollständige Subskala.)

Variablen-name	Item	Pre α	Post α	Fol α
pra5, poa40, foa16	Der Aufbau von Versuchsanordnungen zum Experimentieren fällt mir leicht.	0,729	0,817	0,811
pra2, poa47, foa8	Ich bin gut darin, Experimente durchzuführen.			
pra17, poa37, foa33	Es fällt mir leicht, die zu einem Experiment benötigten Geräte zu bedienen.			

Tabelle 56: Items und Reliabilität (Cronbachs α) der FSWEx.V2-Subskala Auswertung im Pre-, Post- und Follow-up-Test der Nebenstudie (Angabe sind nur die Items, um die der FSWEx.V2 den FSWEx.V1 ergänzt. Die Reliabilitäten beziehen sich auf die vollständige Subskala.)

Variablen-name	Item	Pre α	Post α	Fol α
pra3, poa1, foa21	Es ist einfach, Ergebnisse aus einem Experiment auszuwerten.	0,786	0,825	0,837
pra24, poa16, foa29	Bei einem nicht planmäßig verlaufenen Experiment fällt es mir leicht, mögliche Fehlerquellen zu finden.			
pra33, poa49, foa3	Es fällt mir leicht, aus experimentellen Messdaten Grafiken zu erstellen.			

2.2.3.4 Was wurde in der Lehrkräfte-Befragung erhoben?

Im folgenden Abschnitt werden die in den Fragebögen der Lehrer eingesetzten Messinstrumente vorgestellt. Eine Übersicht der verwendeten Messinstrumente liefert Tabelle 57. Als personenbezogene Daten wurde ausschließlich Geschlecht und Alter der Lehrer erfragt. Die Befragung der Lehrer wurde lediglich zu einem Messzeitpunkt durchgeführt.

Tabelle 57: Übersicht über die in der Lehrerbefragung eingesetzten Messinstrumente (Die Befragung wurde nur an einem Messzeitpunkt durchgeführt)

Messinstrumente der Lehrerbefragung
<ul style="list-style-type: none"> • Personenbezogene Daten (Geschlecht, Alter) • Frage zu Problemen beim Experimentieren in der Schule • Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung in den Subskalen Planung, Durchführung und Auswertung (FSWEx.V1) • Skala zur Sicherheit beim Experimentieren • Skala wertbezogenes Interesse

Erhebung der Probleme beim Experimentieren in der Schule aus Lehrersicht

Das BeLL Bio möchte Schülern die Möglichkeit geben, in der authentischen Lernumgebung Labor solche Experimente durchzuführen, die in der Schule so nicht möglich sind. Welche Probleme Lehrer beim Experimentieren in der Schule sehen, sollte mit der folgenden Frage geklärt werden (Abb. 29).

Welche Probleme sehen Sie, wenn Sie im Biologieunterricht ihrer Schule experimentieren möchten?

- ☐ Ich sehe keine Probleme beim Experimentieren in meinem Biologieunterricht.
- ☐ Der straffe Lehrplan lässt mir wenig Zeit für Experimente.
- ☐ Die materielle Ausstattung des Unterrichtsraumes lässt kaum Experimente zu.
- ☐ Die Verbrauchsmaterialien (z.B. Chemikalien, Einweggefäße) sind vom Schuletat kaum zu finanzieren.
- ☐ Einige Experimente führe ich deshalb ungerne durch, weil ich mich bei diesen Experimenten unsicher fühle.
- ☐ _____

Abbildung 29: Frage zur Erfassung der Probleme des Experimentierens in der Schule im Lehrerfragebogen

Experimentiersicherheit

Als mögliches Problem schulischen Experimentierens wurde in der o.g. Frage (Abb. 29) die Selbstsicherheit beziehungsweise Unsicherheit beim Experimentieren erfragt. Wie stark die Selbstsicherheit tatsächlich ausgeprägt ist, wurde mit den Items der Skala Experimentiersicherheit in Tabelle 58 über eine fünfstufige Likert-Skala („stimmt gar nicht“ bis „stimmt völlig“) erfasst.

Tabelle 58: Items und Reliabilität (Cronbachs α) der Messdimension Experimentiersicherheit im Lehrerfragebogen (Das grau hinterlegte Item wurde im Rahmen der Reliabilitätsoptimierung aus der Auswertung ausgeschlossen und wird im angegebenen Wert für Cronbachs α nicht berücksichtigt, *Items wurden umkodiert)

Variablen- name	Item	α
Lea2*	Beim Durchführen von Experimenten im Unterricht fühle ich mich unsicher.	0,732
Lea9	Im Unterricht traue ich mich auch an neue Versuche heran.	
Lea16	Mir macht es Spaß, im Unterricht neue Experimente auszuprobieren.	
Items neu entwickelt		

Fragebogen zur subjektiv wahrgenommenen Experimentierkompetenz

Die zuvor beschriebene Skala zur Experimentiersicherheit erfasst die subjektive Höhe der empfundenen Selbstsicherheit. Sie misst nicht, in welcher Teilkompetenz des Experimentierens die Lehrer eine hohe bzw. geringe Selbstwirksamkeit bekunden. Zu diesem Zweck wurde der bereits im Rahmen der Hauptstudie (vgl. Punkt 2.2.3.2.3.1) vorgestellte FSWEx.V1 eingesetzt. Tabelle 59 gibt einen Überblick über die in der Lehrerbefragung erzielten Reliabilitäten.

Tabelle 59: Reliabilitäten (Cronbachs α) der FSWE.V1-Subskalen in der Lehrerbefragung (Es wurden keine Items zur Reliabilitätsoptimierung ausgeschlossen.)

FSWE.V1 -Subskala	Anzahl der Items	Reliabilität
Planung	3	0,726
Durchführung	3	0,700
Auswertung	3	0,626

Interesse – Wertebezogene Komponente

Um mögliche Differenzen im wahrgenommenen Nutzen von Schülerlaborkursen zwischen selbstsicheren und unsicheren Lehrern erfassen zu können, wurde die folgende Skala eingesetzt (Tab. 60). Sie orientiert sich an der Skala zur Erfassung der wertebezogenen Komponente des aktuellen Interesses von Engeln (2004). Die Items wurden allerdings umformuliert.

Tabelle 60: Items und Reliabilität (Cronbachs α) der Messdimension Interesse – Wertebezogene Komponente im Lehrerfragebogen

Variablen- name	Item	α
Lea3	Die Inhalte der Laborkurse sind für mich als Lehrer wichtig.	0,828
Lea13	Die Inhalte der Laborkurse können mir als Lehrer nützlich sein.	
Lea18	Die Inhalte der Laborkurse sind für mich als Lehrer sinnvoll.	
Items in Anlehnung an Engeln (2004)		

2.2.4 Überblick über die evaluierten Laborkurse

Nachdem in den vorangegangenen Abschnitten die in den Studien eingesetzten Messinstrumente präsentiert wurden, werden im Anschluss die im Rahmen der Haupt- und Nebenstudie evaluierten, im Schülerlabor BeLL Bio an der Bergischen Universität Wuppertal durchgeführten, Laborkurse vorgestellt. Wie Engeln (2004) feststellt, sind die in Schülerlaboren durchgeführten Experimente so anzulegen, dass die Kursteilnehmer nicht überfordert werden. Die Vermeidung von Überforderung bei gleichzeitiger Herausforderung kann aber nur gelingen, wenn die Lernvoraussetzungen der Schüler berücksichtigt werden.

2.2.4.1 Lernvoraussetzungen der Schüler

Wie bereits im Rahmen der Stichprobenbeschreibung dargestellt, handelt es sich bei den Teilnehmern der Haupt- und Nebenstudie um Schüler der Sekundarstufe II. Als zentrale Methode der naturwissenschaftlichen Erkenntnisgewinnung wird dem Experiment sowohl in den Richtlinien und Lehrplänen der Sekundarstufe I als auch der Sekundarstufe II eine zentrale Bedeutung beigemessen (vgl. Punkt 1.2.4.2). Grundlegende Erfahrungen der Schüler im praktischen Experimentieren sollten

daher vorausgesetzt werden können. Wie einschlägige Studien berichten, ist dem allerdings nicht so. Engeln (2004) stellt beispielsweise für den Physikunterricht der Jahrgangsstufen neun und zehn fest, dass Schülerexperimente deutlich seltener durchgeführt werden als Demonstrationsexperimente durch die Lehrperson. Ähnliches berichtet Scharfenberg (2005). In den Kursen des BeLL Bio ist daher mit einer, in Bezug auf die praktische Experimentiererfahrung, sehr heterogenen Schülerschaft zu rechnen.

Damit Schüler die Lernumgebung BeLL Bio als angemessene Herausforderung betrachten, ist es nicht nur notwendig, ihre praktischen Fähigkeiten, sondern auch ihre theoretischen Vorerfahrungen zu berücksichtigen. Die enge curriculare Anbindung der BeLL Bio-Kurse an das Unterrichtsfach Biologie hilft dabei, diese einzuschätzen. Tabelle 61 bietet einen Überblick über die Anknüpfung der evaluierten BeLL Bio-Kurse Einzeller, LDH und Genetischer Fingerabdruck an die Richtlinien und Lehrpläne der Sekundarstufe II Gym/Ges NRW im Fach Biologie.

Tabelle 61: Anbindung der BeLL Bio-Kurse an das Curriculum der Sekundarstufe II im Fach Biologie (vgl. MSWWF NRW 1999)

Genetischer Fingerabdruck		Anknüpfungspunkte an die Richtlinien und Lehrpläne der Sekundarstufe II Gym/Ges NRW Biologie (1999)
Der genetische Fingerabdruck – Dem Täter auf der Spur Inhalt: <ul style="list-style-type: none"> • Durchführung einer DNA-Extraktion aus Mundschleimhautzelle • Amplifikation eines VNTR-Locus mittels PCR • Auftrennung der amplifizierten PCR-Produkte mittels Gelelektrophorese 		Jahrgangsstufe 12/13 (nach G8: Q1/Q2) Unter dem Leitthema <i>Genetische und entwicklungsbiologische Grundlagen von Lebensprozessen</i> werden hier obligatorisch unter anderem die Themenfelder <i>Molekulare Grundlagen der Vererbung und Entwicklungssteuerung</i> sowie <i>Angewandte Genetik</i> behandelt. Obligatorisch werden in diesen Themenfeldern unter anderem die DNA als Träger der Erbinformation, die DNA-Replikation, die Werkzeuge und Verfahrensschritte der Gentechnik und als obligatorische Fachmethode die PCR behandelt. Grundlagen der Reproduktion und Vererbung werden bereits in der Sekundarstufe I besprochen.
Außer Atem – Sportliche Enzyme Inhalt: <ul style="list-style-type: none"> • Anaerober Stoffwechsel (Milchsäuregärung) • Photometrische Bestimmung der Pyruvatkonzentration in zwei unbekannten Pyruvatproben 		Jahrgangsstufe 11 Unter dem Leitthema <i>Physiologie: Struktur – Funktion – Wechselwirkung</i> wird hier obligatorisch das Themenfeld <i>Biokatalyse</i> behandelt. Inhaltlich werden hier verpflichtend der molekulare Bau und die Wirkungsweise von Enzymen besprochen. Als mögliche zu behandelnde Fachmethode wird die Photometrie vorgeschlagen. Weiterhin wird in dieser Jahrgangsstufe das Themenfeld <i>Betriebsstoffwechsel und Energieumsatz</i> behandelt. Obligatorischer Inhalt ist hier unter anderem der anaerobe Abbau von Glucose und Gärung. In der Sekundarstufe I wird zudem die

Genetischer Fingerabdruck	Anknüpfungspunkte an die Richtlinien und Lehrpläne der Sekundarstufe II Gym/Ges NRW Biologie (1999)
	<p>Wirkungsweise von Enzymen modellhaft vorgestellt (Schlüssel-Schloss-Prinzip).</p> <p>Die Bedeutung von ATP als Energieträger wird erst in den Jahrgangsstufen 12/13 im Themenfeld <i>Molekulare und cytologische Grundlagen</i> behandelt.</p>
<p>Kleine Einzeller ganz groß</p> <p>Inhalt:</p> <p>Versuche zur/zum</p> <ul style="list-style-type: none"> • Zellkernfärbung (<i>Tetrahymena thermophila</i>) • Phagocytose (<i>Tetrahymena thermophila</i>) • Konjugation (<i>Tetrahymena thermophila</i>) • Einfluss von Tabakrauch auf die Cilienbewegung von <i>Euglena gracilis</i> 	<p>Jahrgangsstufe 11</p> <p>Unter dem Leitthema <i>Physiologie: Struktur – Funktion – Wechselwirkung</i> wird hier obligatorisch das Themenfeld <i>Zelle, Gewebe, Organismus</i> behandelt. Inhaltlich werden u.a. das lichtmikroskopische Bild der Zelle und der Zellkern besprochen. Verpflichtende Methoden sind die Verwendung des Lichtmikroskops, das Anfertigen, Zeichnen und Auswerten von Präparaten.</p> <p>Einem weiteren obligatorischen Themenfeld gehören <i>Molekulare Grundlagen, Kompartimentierung, Transport</i> an. Hier werden u.a. der Bau und die Funktion von Biomembranen sowie die Transportvorgänge an Membranen behandelt.</p> <p>Im Lehrplan der Sekundarstufe II wird vorausgesetzt, dass in der Sekundarstufe I bereits lichtmikroskopisch sichtbare Grundstrukturen der Zelle besprochen und mikroskopische Übungen durchgeführt wurden.</p> <p>Auch Grundlagen der Reproduktion und Vererbung werden bereits in der Sekundarstufe I behandelt.</p>

Die Schüler besuchen die Laborveranstaltungen Genetischer Fingerabdruck und Einzeller in der Regel erst zum Ende einer entsprechenden Unterrichtseinheit. Die Laborkursinhalte können daher gezielt das theoretische Vorwissen der Schüler aufgreifen und anhand praktischer Erfahrungen vertiefen. Gerade für den Laborkurs Genetischer Fingerabdruck gilt zu beachten, dass dieser Kurs sowohl von Leistungs-, als auch von Grundkursen besucht wird. Bedingt durch die geringere Wochenstundenzahl der Grundkurse ist hier mit einem weniger tiefgreifenden Vorwissen zu rechnen, als in den Leistungskursen. Ein Besuch des LDH-Kurses wird von vielen Lehrern im Rahmen der Behandlung des Themas Biokatalyse gewünscht. Häufig war zu diesem Zeitpunkt das Thema anaerober Stoffwechsel noch nicht Gegenstand des Unterrichts, so dass entsprechende theoretische Grundlagen fehlen. Einen grundlegenden Überblick über den anaeroben Stoffwechsel gilt es daher im Laborkurs zu vermitteln.

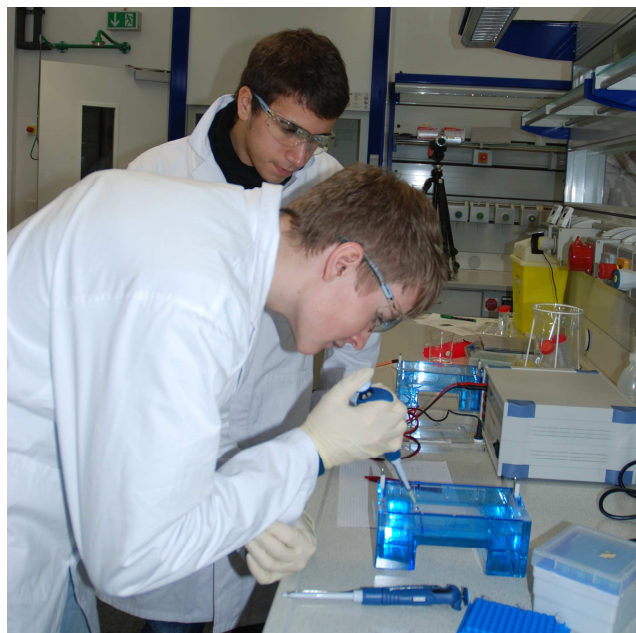
Neben ihren fachpraktischen Fähigkeiten und ihrem theoretischen Vorwissen bringen Schüler ihre persönlichen Alltagserfahrungen mit ins Schülerlabor. Insbesondere der Genetische Fingerabdruck ist häufig aus einschlägigen TV-Serien geläufig. Ein tiefergehender Einblick in die fachliche Methodik sowie den tatsächlichen Beweischarakter des Genetischen Fingerabdrucks wird hier allerdings kaum vermittelt. Auch verleiten die Kürze der Darstellung und die oftmals rasanten Schnittfolgen zu der Annahme, innerhalb weniger Minuten von der Tatortspur zum fertigen Fingerprint zu gelangen. Diese Schülervorstellungen gilt es im BeLL Bio-Kurs aufzugreifen und in fachlich korrektere Vorstellungen zu überführen.

Ein Überblick über Ziele und Inhalte der drei evaluierten Laborkurse wird in den folgenden Abschnitten gegeben.

2.2.4.2 Laborkurs: „Der Genetische Fingerabdruck – Dem Täter auf der Spur“

(Kurskürzel: Genetischer Fingerabdruck)

Wie der Titel des Kurses bereits verrät, werden in diesem Laborkurs zentrale Methoden der Molekularbiologie im Kontext der Kriminalistik angewandt. Aufgabe der Schüler ist es, einen Tatverdächtigen anhand seines genetischen Fingerabdrucks zu überführen, beziehungsweise einen zweiten Tatverdächtigen zu entlasten. Dies gelingt ihnen mittels DNA-Extraktion aus Mundschleimhautzellen sowie einer PCR mit anschließender Gelelektrophorese (Abb. 30).



**Abbildung 30: Schüler beim Auftragen von DNA-Proben auf ein Agarosegel
im BeLL Bio-Kurs Genetischer Fingerabdruck**

Erwartete Lernziele im Laborkurs Genetischer Fingerabdruck

Die Lernziele des Laborkurses werden in Tabelle 62 zusammengefasst.

Tabelle 62: Lernziele des BeLL Bio-Kurses Genetischer Fingerabdruck

Grobziel	<ul style="list-style-type: none"> Die Schüler sollen die zentralen laborpraktischen Schritte kennenlernen, die zur Erstellung eines genetischen Fingerabdrucks notwendig sind.
Feinziele kognitiv	<ul style="list-style-type: none"> Die Schüler sollen die Methoden der DNA-Extraktion, PCR und Gelelektrophorese erklären können. Die Schüler sollen die zur Beschreibung der Labormethoden notwendigen biowissenschaftlichen Fachbegriffe korrekt anwenden können.
Feinziele affektiv	<ul style="list-style-type: none"> Durch die aktive Auseinandersetzung mit den Fachinhalten sollen die Schüler ein hohes aktuelles Interesse generieren. Indem sich die Schüler im Laborkurs als kompetent erleben, sollen sie ihr biologisches Fähigkeitsselbstkonzept und ihre experimentbezogene Selbstwirksamkeit steigern.
Feinziele psychomotorisch	<ul style="list-style-type: none"> Die Schüler sollen Mikroliterpipetten bedienen können. Die Schüler sollen eine Gelelektrophorese durchführen können.

Ablauf des Laborkurses Genetischer Fingerabdruck

Der Laborkurs Genetischer Fingerabdruck beginnt um 09:00 Uhr und dauert etwa sechseinhalb Zeitstunden. Treffpunkt ist der Haupteingang der Bergischen Universität Wuppertal. Hier nimmt der Laborkursleiter die Schüler in Empfang. Nach Vorstellung und Begrüßung werden die Schüler in den Seminarraum des nur etwa zwei Gehminuten entfernten BeLL Bio geführt. Dort findet auch die Vorbesprechung für die Laborarbeit statt. Zu Beginn der Power-Point gestützten Vorbesprechung werden die Schüler vom Laborkursleiter mit einem fiktiven Zeitungsartikel konfrontiert, dessen Titel lautet: „Schlächter schlägt wieder zu – Polizei fasst zwei Tatverdächtige“. Der Artikel schildert, wie ein Mann vor seiner Haustür von einer unbekannten Person niedergeschlagen wird. Unmittelbar nach der Tat wird das Opfer von seiner Frau gefunden. Diese sieht in einiger Entfernung noch eine grün gekleidete Person in Richtung einer Parkanlage davonlaufen. Eine Luftaufnahme des Tatortes wird den Schülern präsentiert. Im Rahmen eines Unterrichtsgespräches motiviert der Laborkursleiter die Kursteilnehmer dazu sich zu äußern, wie sie an Stelle der Ehefrau des schwer verletzten Opfers handeln würden. Zu erwarten ist, dass die Kursteilnehmer schnell auf die Verständigung des Notarztes sowie der Polizei kommen. Als notwendige Ermittlungsmaßnahmen der Polizei werden die Spurensuche am Tatort sowie die Fahndung nach dem Täter genannt. Die am Tatort gefundenen subjektiven (Zeugenaussage der Ehefrau des Opfers) und objektiven Spuren (Speichel des Täters auf der Kleidung des Opfers, Tatwerkzeug) sowie die in Tatortnähe angetroffenen Tatverdächtigen A und

B werden vom Laborkursleiter vorgestellt. Weiterhin erhalten die Schüler die Information, dass die Polizei bereits Speichelproben von beiden Tatverdächtigen entnommen hat. Der Laborkursleiter fordert die Kursteilnehmer auf, Vermutungen über die Bedeutung der Speichelproben für die Aufklärung der Straftat zu äußern. Da in der unterrichtlichen Vorbereitung bereits das Versuchsskript besprochen wurde, sollten schnell die Zellen der Mundschleimhaut als Quelle der Tatverdächtigen-DNA identifiziert werden. Den Schülern wird mitgeteilt, dass von der Polizei bereits die DNA der Speichelproben (Probe A=Tatverdächtiger A, Probe B=Tatverdächtiger B, Probe T=Täter-DNA der Speichelprobe auf der Kleidung des Opfers) isoliert wurde und noch die Methoden der PCR und Gelelektrophorese zur Erstellung des genetischen Fingerabdrucks ausstehen. Die Schüler erhalten zum Ende der Vorbesprechung die Aufgabe, den Straftäter anhand seines genetischen Fingerabdrucks zu überführen. Nach einer Sicherheitsbelehrung, einer Einweisung in die Bedienung der Mikroliterpipetten und dem Austeilen der Laborkittel beginnt die praktische Arbeit im Labor.

Wie bei den anderen evaluierten Laborkursen, so arbeiten die Schüler auch in diesem Kurs in Zweier-Arbeitsgruppen. Der Kurs wird neben dem Laborkursleiter von zwei bis – in Abhängigkeit von der Anzahl der Kursteilnehmer – drei Betreuern begleitet. Da der praktische Umgang mit Mikroliterpipetten den meisten Kursteilnehmern noch nicht vertraut ist, beginnt der Laborkurs mit einer Pipettierübung. Aufgabe ist es hier, eine der drei zur Verfügung stehenden Mikroliterpipetten sowie ein für die Pipette geeignetes Volumen und passende Pipettenspitzen auszuwählen und Wasser von einem Becherglas in ein anderes zu pipettieren. Im Anschluss daran erhält jede Arbeitsgruppe eine DNA-Probe (A, B oder T) sowie einen PCR-Mix, um die PCR zur Amplifizierung des D1S80-Abschnittes auf dem humanen Chromosom 1 anzusetzen. Die Laufzeit der PCR wird von den Schülern genutzt, um Zellen aus der eigenen Mundschleimhaut zu gewinnen und DNA aus diesen zu isolieren. Allerdings wird mit der selbst extrahierten DNA nicht weitergearbeitet. Im Laborkurs wird bewusst nicht die im Laboralltag übliche Reihenfolge DNA-Isolierung – PCR – Gelelektrophorese gewählt, da sich eine Verlängerung des ohnehin schon relativ zeitintensiven Laborkurses um etwa neunzig Minuten ergeben würde. Wie Testläufe der Laborveranstaltung zeigten, nimmt die Bereitschaft der Schüler zur Mitarbeit in einem derart langen Kurs erheblich ab. Diese Testläufe sind nicht Bestandteil der Evaluierung.

Die PCR-Produkte der Proben A, B bzw. T werden schließlich mittels Gelelektrophorese aufgetrennt. Die Laufzeit des Gels wird für eine einstündige Mittagspause genutzt. Nach der Mittagspause folgt die Betrachtung und Diskussion der Gelbanden am UV-Transilluminator. Durch Vergleich der Gelbanden der Tatverdächtigen A und B mit den Banden der Tatort-DNA wird es den Schülern möglich, einen der beiden Tatverdächtigen der Tat zu überführen. In der PowerPoint-gestützten Nachbesprechung werden ein exemplarisches Gelbild und mögliche Fehlerquellen im Versuch

diskutiert sowie die Beweiskraft eines einzelnen untersuchten DNA-Abschnittes erörtert. Weitere Anwendungsmöglichkeiten des genetischen Fingerabdrucks werden im Kontext verschiedener Berufsfelder diskutiert.

Tabellarische Planung des Laborkurses Genetischer Fingerabdruck

In der nachfolgenden Tabelle 63 wird ein Überblick über den Ablauf der Laborveranstaltung Genetischer Fingerabdruck gegeben.

Tabelle 63: Tabellarische Planung des Laborkurses Genetischer Fingerabdruck (BUW=Bergische Universität Wuppertal, LP=Lehrperson, PPP=PowerPoint-Präsentation, SuS=Schülerinnen und Schüler)

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
09:00h	Begrüßung (Haupteingang der BUW)	Schülergruppen werden vom Laborkursleiter am Haupteingang der BUW in Empfang genommen. Anschließend wird gemeinsam der Seminarraum des BeLL Bio aufgesucht.	Vorstellung des Laborkursleiters. Begrüßung an der Universität zur Betonung der authentischen Lernumgebung. Treffen von Absprachen mit der begleitenden LP bezüglich verspäteter SuS. Erste Einschätzung der Aufmerksamkeit und Motivation der SuS möglich.	---
09:15h	Einstieg (Seminarraum BeLL Bio)	Anhand eines fiktiven Zeitungsartikels werden die SuS mit einem Mordfall konfrontiert. Die SuS werden aufgefordert, Ideen zur Lösung des Falls zu äußern. Ihre Ideen führen zur Ermittlung von zwei Tatverdächtigen und zum Auffinden einer Speichelspur des Täters am Tatort. Laborkursleiter lässt SuS erklären, wie mit der Speichelspur vorzugehen ist, um den Täter zu überführen.	Zeitungsartikel liefert Problemstellung. Gelenktes Unterrichtsgespräch. Laborkursleiter gibt Anregungen durch gezielte Nachfragen.	PPP, Beamer
09:45h	Sicherheitsbelehrung + Einweisung in	Laborkursleiter führt eine verpflichtende Sicherheitsbelehrung durch, die	Gewährleistet sicheres Arbeiten im Labor und sicheren Umgang mit den im	PPP, Beamer, Mikroliter-

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
	Mikroliter-pipetten + Verteilung der Arbeitsplätze (Seminarraum + Laborraum BeLL Bio)	von den SuS zu unterzeichnen ist. Laborkursleiter führt Mikroliterpipetten vor. Verteilen der Laborkittel und Belegung der Arbeitsplätze	Labor zu verwendenden Arbeitsgeräten. Persönliche Verantwortung für einen Arbeitsplatz erhöht Sorgfalt.	pipetten
10:00h	Erarbeitung I (Laborraum BeLL Bio)	SuS führen Pipettierübung durch.	SuS eignen sich selbstständig in 2er-Gruppen eine wichtige Voraussetzung für die erfolgreiche Durchführung des Versuches an.	Versuchsskript, Mikroliter-pipetten, Sicherheits-ausstattung (Kittel, Schutzbrillen, Handschuhe)
10:10h	Erarbeitung II (Laborraum BeLL Bio)	SuS setzen PCR mit bereits extrahierter DNA der Tatverdächtigen/des Täters an. SuS nutzen die Dauer der PCR für die Extraktion der DNA aus ihren eigenen Mundschleimhautzellen. Hier setzen sie wiederholt die Methode des Zentrifugierens ein. Schnelle SuS-Gruppen gießen bereits die Gele. DNA-Extraktion und Gele sind in etwa dann fertig, wenn die PCR beendet ist. SuS können ihre PCR-Produkte auf die Gele auftragen.	Versuchsskript-geleitete Arbeit in 2er-Gruppen gewährleistet hohe Aktivität aller SuS. Nutzung der PCR-Dauer für die DNA-Extraktion gewährleistet die Vermeidung langer Wartezeiten. Beladen der Zentrifugen und Gele erfordert Absprache der SuS-Gruppen untereinander. Nutzung eines „sicheren“ UV-Farbstoffes gewährleistet hohe Arbeitssicherheit.	Versuchsskript, Mikroliter-pipetten, Reaktions-gefäße, Zentrifugen, Thermocycler, Gelkammern, Sicherheits-ausstattung (Kittel, Schutzbrillen, Handschuhe)
13:00h	Mittagspause	Gelelektrophorese läuft. SuS besuchen selbstständig Mensa oder Cafeteria.	Die Dauer der Gelelektrophorese entspricht der Länge der Mittagspause.	---
14:00h	Ergebnis-sicherung I (Laborraum BeLL Bio)	Entnahme der Gele aus den Gelkammern und Visualisierung der DNA-Banden auf dem UV-Transilluminator.	SuS werden vom Laborkursleiter auf Sicherheitsaspekte hingewiesen. SuS sehen das Ergebnis ihrer Arbeit als DNA- Banden. SuS können Täter überführen.	UV-Trans-illuminator Agarosegel
14:30h	Ergebnis-sicherung II	Bandenmuster wird an einem exemplarischen Geldbild	Gelenktes Unterrichtsgespräch.	PPP, Beamer

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
	(Seminarraum BeLL Bio)	besprochen. Weitere Anwendungsmöglichkeiten des genetischen Fingerabdruckes werden im Kontext verschiedener Berufsfelder angesprochen.	Sollte die PCR nicht funktioniert haben, so ermöglicht ein vorbereitetes Gelbild dennoch eine Ergebnissicherung.	
15:15h-15:35h	Posttest (Seminarraum BeLL Bio)	SuS füllen Fragebogen aus.	Bestandteil des Pre-, Post-, Follow-up-Testdesigns.	Fragebogen

2.2.4.3 Laborkurs: „Außer Atem – Auf der Spur sportlicher Enzyme“ (Kurskürzel: LDH)

Der vierstündige Laborkurs wurde gemeinsam mit Aepfelbach (2011) entwickelt. Versuchsskript und Kurs wurden allerdings im Rahmen der vorliegenden Studie überarbeitet. Inhaltlicher Schwerpunkt des Kurses ist die Durchführung eines optischen Tests zur Bestimmung von Substratkonzentrationen. Eingebettet in einen sportlichen Kontext ist es Aufgabe der Schüler, die Pyruvatkonzentration in zwei unbekannten künstlichen Blutproben herauszufinden. Dies gelingt ihnen unter Anwendung der Photometrie (Abb. 31).



Abbildung 31: Schülerinnen bei der Arbeit am Photometer im BeLL Bio-Kurs LDH

Lernziele des Laborkurses LDH

In Tabelle 64 werden die Lernziele des Laborkurses LDH zusammengefasst.

Tabelle 64: Lernziele des BeLL Bio-Kurses LDH

Grobziel	<ul style="list-style-type: none"> Die Schüler sollen die laborpraktischen Schritte erklären können, die zur photometrischen Bestimmung der Pyruvatkonzentration notwendig sind.
Feinziele kognitiv	<ul style="list-style-type: none"> Die Schüler sollen die Methode der Photometrie erklären können. Die Schüler sollen die Bedeutung der Laktatdehydrogenase für den Pyruvatnachweis erläutern können. Die Schüler sollen die Bedeutung von NADH für den photometrischen Pyruvatnachweis erläutern können. Die Schüler sollen das Lambert-Beer'sche-Gesetz zur Berechnung von Konzentrationen anwenden können. Die Schüler sollen die zur Beschreibung der Labormethoden notwendigen biowissenschaftlichen Fachbegriffe korrekt anwenden können.
Feinziele affektiv	<ul style="list-style-type: none"> Durch die aktive Auseinandersetzung mit den Fachinhalten, sollen die Schüler ein hohes aktuelles Interesse generieren. Indem sich die Schüler im Laborkurs als kompetent erleben, sollen sie ihr biologisches Fähigkeitsselbstkonzept und ihre experimentbezogene Selbstwirksamkeit steigern.
Feinziele psychomotorisch	<ul style="list-style-type: none"> Die Schüler sollen Mikroliterpipetten bedienen können. Die Schüler sollen Absorptionsmessungen mit dem Photometer durchführen können.

Ablauf des Laborkurses LDH

Die Gesamtdauer des Kurses „Außer Atem – Auf der Spur sportlicher Enzyme“ beträgt vier Zeitstunden. In der Regel werden die Biologiekurse vom Laborkursleiter gegen 09:00h am Haupteingang der Bergischen Universität Wuppertal in Empfang genommen, um im Anschluss den Seminarraum des Schülerlabors BeLL Bio aufzusuchen. Hier findet auch die Power-Point-gestützte Vorbesprechung des Laborkurses statt, zu deren Beginn der Laborkursleiter den Schülern ein Bild von zwei Personen präsentiert, die nach sportlicher Anstrengung außer Atem geraten sind. Die Schüler werden animiert, Erklärungen für dieses allgemein bekannte Problem zu liefern. Erwartete Antworten sind, dass die Personen anstrengungsbedingt viel Sauerstoff verbraucht haben und der Sauerstoff bei anhaltender Belastung nicht schnell genug nachgeliefert werden kann. Biologiekursen, die im Unterricht bereits aeroben und anaeroben Stoffwechsel besprochen haben, sollte es möglich sein, den Ort und die Bedeutung des Sauerstoffverbrauchs in den Zellen näher zu lokalisieren. Im Rahmen eines gelenkten Unterrichtsgesprächs wird den Schülern, gestützt durch die Abbildung eines vereinfachten Schemas des Glucoseabbaus in der Zelle, die Problematik des Sauerstoffmangels und

die Notwendigkeit eines anaeroben Stoffwechselweges vermittelt. Als zentrales Enzym des anaeroben Stoffwechsels wird die Laktatdehydrogenase (LDH) sowie die von ihr katalysierte Umsetzung von Pyruvat zu Laktat unter NADH-Verbrauch vorgestellt. An dieser Stelle werden die Schüler gefragt, wie sie vorgehen würden, um die Pyruvatkonzentration in einer Flüssigkeit zu bestimmen. Da das Versuchsskript bereits im Biologieunterricht besprochen wurde, sollten die Schüler die Photometrie als mögliche Methode in Erwägung ziehen. Der Aufbau eines Photometers sowie das Lambert-Beer'sche-Gesetz werden besprochen. Weiterhin wird die Problematik diskutiert, Pyruvat im Wellenlängenbereich UV/VIS nicht direkt messen zu können, so dass der indirekte Weg der Konzentrationsbestimmung über den NADH-Abbau gewählt werden muss (vgl. Abschnitt 1.1.2). Nach einer Sicherheitsbelehrung, einer Einweisung in die Bedienung der Mikroliterpipetten und dem Austeilen der Laborkittel beginnt in diesem Kurs die praktische Arbeit im Labor. Aufgabe der Schüler ist es – geleitet durch das Versuchsskript und mit Unterstützung der anwesenden Betreuer – die Pyruvatkonzentration in zwei künstlichen Blutproben zu bestimmen.

Wie im Laborkurs Genetischer Fingerabdruck, so beginnt auch hier die praktische Arbeit mit einer Pipettierübung. Auch dieser Kurs wird neben dem Laborkursleiter von zwei bis drei Betreuern begleitet. Im Laborraum arbeiten die Schüler – angeleitet durch das Versuchsskript – in Zweierteams zusammen. Der Raum ist mit fünf Photometern und Laptops ausgestattet. Längere Wartezeiten an den Photometern können vermieden werden, da jede Gruppe den NADH-Abbau zunächst nur in einer der beiden vorliegenden Proben photometrisch verfolgen soll. Die Messwerte werden in tabellarischer Form im Versuchsskript notiert. Anschließend werden die Messwerte in eine vorbereitete Eingabemaske des Tabellenkalkulationsprogramms OpenOffice CALC übertragen. Aus den Daten wird eine Kurve generiert, die die Abnahme der NADH-Konzentration über die Zeit repräsentiert. Nachdem die Daten gespeichert wurden, erfolgt die Auswertung der zweiten Blutprobe nach dem gleichen Prinzip. Danach werden die beiden Kurven ausgedruckt und können am Ende von den Schülern als gesichertes Ergebnis mitgenommen werden. Nach Abschluss der Experimentierphase steht dann ausreichend Zeit zur Verfügung, um aus den Aufzeichnungen mit Hilfe des Lambert-Beer'schen-Gesetzes die Pyruvatkonzentration berechnen zu können.

In der Nachbesprechung werden exemplarische NADH-Abbaukurven der Schüler besprochen. Eine vorbereitete PowerPoint-Präsentation erlaubt es, an einem Beispiel die Berechnung der Pyruvatkonzentration aus dem NADH-Abbau nachzuvollziehen. Abschließend werden die in den ausgegebenen künstlichen Blutproben vorhandenen NADH-Konzentrationen bekannt gegeben, so dass die Schüler diese mit ihren eigenen Ergebnissen abgleichen können. Abschließend werden weitere Einsatzbereiche der Photometrie in verschiedenen Berufsfeldern werden besprochen.

Tabellarische Darstellung des Laborkurses LDH

In der nachfolgenden Tabelle 65 wird ein Überblick über den Ablauf der Laborveranstaltung LDH gegeben.

Tabelle 65: Tabellarische Planung des Laborkurses LDH (BUW=Bergische Universität Wuppertal, LDH=Laktatdehydrogenase, LP=Lehrperson, PPP=PowerPoint-Präsentation, SuS=Schülerinnen und Schüler)

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
09:00h	Begrüßung (Haupteingang der BUW)	Schülergruppen werden vom Laborkursleiter am Haupteingang der BUW in Empfang genommen. Anschließend wird gemeinsam der Seminarraum des BeLL Bio aufgesucht.	Vorstellung des Laborkursleiters. Begrüßung an der Universität zur Betonung der authentischen Lernumgebung. Treffen von Absprachen mit der begleitenden LP bezüglich verspäteter SuS. Erste Einschätzung der Aufmerksamkeit und Motivation der SuS möglich.	---
09:15h	Einstieg (Seminarraum BeLL Bio)	Den SuS wird ein Bild von zwei Personen gezeigt, die nach sportlicher Betätigung außer Atem geraten. SuS sollen die biologischen Grundlagen für das im Bild beobachtete Phänomen erläutern. Laborkursleiter diskutiert mit SuS die Bedeutung der LDH im anaeroben Stoffwechsel. SuS erhalten die Aufgabe, photometrisch die Pyruvatkonzentration in zwei unbekannten künstlichen Blutproben zu bestimmen. Der Aufbau eines Photometers und das Lambert-Beersche-Gesetz werden besprochen.	Bild liefert Problemstellung. Gelenktes Unterrichtsgespräch. Laborkursleiter gibt Anregungen durch gezieltes Nachfragen. Diskussion des anaeroben Stoffwechsels kann in Abhängigkeit vom Vorwissen der SuS lehrer- oder schülerzentrierter ausfallen. Bei der Besprechung des Photometers wird Vorwissen aus dem Physikunterricht aufgegriffen (Prisma).	PPP, Beamer
10:00h	Sicherheitsbelehrung + Einweisung in Mikroliterpipetten + Verteilung der Arbeitsplätze (Seminarraum + Laborraum BeLL Bio)	Laborkursleiter führt eine verpflichtende Sicherheitsbelehrung durch, die von den SuS zu unterzeichnen ist. Laborkursleiter führt Mikroliterpipetten vor. Verteilen der Laborkittel und Belegung der Arbeitsplätze.	Gewährleistet sicheres Arbeiten im Labor und sicheren Umgang mit den im Labor zu verwendenden Arbeitsgeräten. Persönliche Verantwortung für einen Arbeitsplatz erhöht Sorgfalt.	PPP, Beamer Mikroliterpipetten

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
10:15h	Erarbeitung I (Laborraum BeLL Bio)	SuS führen Pipettierübung durch.	SuS eignen sich selbstständig in 2er-Gruppen eine wichtige Voraussetzung für die erfolgreiche Durchführung des Versuches an.	Mikroliter-pipetten Versuchsskript Sicherheits-ausstattung (Kittel, Schutzbrillen, Handschuhe)
10:25h	Erarbeitung II (Laborraum BeLL Bio)	SuS führen zunächst die photometrischen Messungen für eine der beiden Pyruvatproben durch. Dann wechseln sie zum Laptop, um ihre Messwerte am Laptop in einem Tabellenkalkulationsprogramm zu visualisieren. Anschließend werden die Messungen für die zweite Probe vorgenommen. Berechnung der Pyruvatkonzentration mit Hilfe des Lambert-Beerschen-Gesetzes.	Versuchsskript-geleitete Arbeit in 2er-Gruppen gewährleistet hohe Aktivität aller SuS. Die abwechselnde Arbeit an Photometer und Laptop vermeidet lange Wartezeiten am Photometer. Umgang mit OpenOffice CALC vermittelt die Bedeutung des Computers als Arbeitsgerät. Verknüpfung von Biologie und Mathematik fördert fächerübergreifendes Denken.	Mikroliter-pipetten Photometer Laptops (OpenOffice CALC) Taschen-rechner Versuchsskript Sicherheits-ausstattung (Kittel, Schutzbrillen, Handschuhe)
12:10 h	Ergebnis-sicherung (Seminarraum des BeLL Bio)	Betreuer druckt SuS die in OpenOffice CALC erstellten NADH-Abbau-Kurven aus. Laborkursleiter bespricht mit SuS exemplarisch die NADH-Kurven, mögliche Fehlerquellen sowie den Weg zur Berechnung der Pyruvatkonzentration. Zum Abschluss gibt der Laborkursleiter die tatsächliche Pyruvatkonzentration der beiden Proben bekannt.	SuS erhalten ihr Ergebnis in Papierform. Gelenktes Unterrichtsgespräch. Laborkursleiter gibt Anregungen durch gezieltes Nachfragen. SuS können ihr berechnetes Ergebnis mit der tatsächlichen Pyruvatkonzentration abgleichen.	Ausdruck der NADH-Abbau-Kurven PPP, Beamer
12:40h- 13:00h	Posttest (Seminarraum BeLL Bio)	SuS füllen Fragebogen aus.	Bestandteil des Pre-, Post-, Follow-up-Testdesigns	Fragebogen

2.2.4.4 Laborkurs „Kleine Einzeller ganz groß“ (Kurskürzel: Einzeller)

Der Kurs Einzeller wurde ursprünglich im Rahmen der BA-Thesis von Kaminska (2010) entwickelt und für die vorliegende Studie nur geringfügig abgewandelt. Im Zentrum der Veranstaltung stehen vier Versuche mit den eukaryotischen Einzellern *Tetrahymena thermophila* und *Euglena gracilis* (Abb.32).



Abbildung 32: Schüler beim Mikroskopieren von *Tetrahymena thermophila* im BeLL Bio-Kurs Einzeller

Lernziele des Laborkurses Einzeller

Tablle 66 gibt einen Überblick über die Lernziele des BeLL Bio-Kurses Einzeller.

Tabelle 66: Lernziele des BeLL Bio-Kurses Einzeller

Grobziel	<ul style="list-style-type: none"> Die Schüler sollen einfache lichtmikroskopische Experimente mit den eukaryotischen Einzellern <i>Tetrahymena thermophila</i> und <i>Euglena gracilis</i> kennenlernen.
Feinziele kognitiv	<ul style="list-style-type: none"> Die Schüler sollen den unter dem Lichtmikroskop beobachteten Ablauf der Phagocytose bei <i>Tetrahymena thermophila</i> beschreiben können. Die Schüler sollen die Bedeutung der Konjugation bei <i>Tetrahymena thermophila</i> erklären können. Die Schüler sollen den Zusammenhang der eingeschränkten Funktion des Flimmerepithels in den Atemwegen und „Raucherhusten“ erklären können. Die Schüler sollen die zur Beschreibung der Labormethoden notwendigen biowissenschaftlichen Fachbegriffe korrekt anwenden können. Die Schüler sollen die Funktion einer Wasserstrahlpumpe erklären können.
Feinziele affektiv	<ul style="list-style-type: none"> Durch die aktive Auseinandersetzung mit den Fachinhalten sollen die Schüler ein hohes aktuelles Interesse generieren. Indem sich die Schüler im Laborkurs als kompetent erleben, sollen sie ihr biologisches Fähigkeitsselbstkonzept und ihre experimentbezogene Selbstwirksamkeit steigern.
Feinziele psychomotorisch	<ul style="list-style-type: none"> Die Schüler sollen Mikroliterpipetten bedienen können. Die Schüler sollen lichtmikroskopische Lebendpräparate von <i>Tetrahymena thermophila</i> und <i>Euglena gracilis</i> herstellen können. Die Schüler sollen eine Zellkernfärbung durchführen können.

Ablauf des Laborkurses Einzeller

Der vierstündige Laborkurs Einzeller beginnt in der Regel gegen 09:00 Uhr. Am Haupteingang der Bergischen Universität Wuppertal nimmt der Laborkursleiter die Biologiekurse in Empfang und begleitet diese dann in den Seminarraum des BeLL Bio. Hier findet eine kurze Vorbesprechung statt, in der Verständnisprobleme der unterrichtlich bereits vorbereiteten Laborversuche erörtert werden. Anschließend folgt eine Belehrung über sicheres Arbeiten im Labor sowie eine Einweisung in die Bedienung der Mikroliterpipetten. Nach Aushändigung der Laborkittel begeben sich die Schüler in den Laborraum.

Der Laborkurs wird neben dem Laborkursleiter von zwei bis drei Hilfskräften betreut. Die Schüler arbeiten wie in den anderen BeLL Bio-Kursen in Zweierteams. Im Laborraum stehen 15 qualitativ hochwertige Lichtmikroskope zur Verfügung, in deren Bedienung die Schüler eingewiesen werden. Anschließend verteilen sich die Kursteilnehmer auf die Arbeitsplätze und führen zunächst eine Pipettierübung durch. Sämtliche Materialien für die Versuche mit *T. thermophila* befinden sich an den Arbeitsplätzen. Lediglich die Mikroskope sind keinem festen Platz zugeordnet. Eine Reihenfolge für die Durchführung der Versuche ist nicht vorgegeben.

Unter jedem der beiden Abzüge des Laborraumes befinden sich die Materialien für den Versuch Smoking *Euglena*, so dass sich die Gruppen bei deren Nutzung abwechseln müssen. Im Rahmen dieses Versuches gilt es, zunächst eine unbehandelte *E. gracilis*-Kultur unter dem Lichtmikroskop zu betrachten. Anschließend werden einige Milliliter dieser Kultur in eine Woulff'sche Flasche überführt, um mittels Wasserstrahlpumpe Rauch einer Zigarette in die Flasche zu saugen. Nach etwa einminütiger Inkubationszeit wird dieser Kultur eine Probe entnommen und erneut unter dem Mikroskop die Motilität von *E. gracilis* überprüft. Die für diesen Versuch notwendigen Materialien (Wasserstrahlpumpe, Woulff'sche Flasche, Plastikschläuche, Zigarette) sind noch nicht zu einem vollständigen Versuchsaufbau zusammengesetzt. Eine detaillierte Anleitung für den Versuchsaufbau finden die Schüler auch in ihrem Versuchsskript nicht. Lediglich das Ziel des Versuchs, nämlich *E. gracilis* Zigarettenrauch auszusetzen, ist hier formuliert. Bevor der Versuch durchgeführt werden darf, sollen die Schüler zunächst den Versuchsaufbau mittels bereitgestellter Pappsablonen der Versuchsmaterialien zusammensetzen und ihren Aufbau mit einem Kursbetreuer absprechen.

Alle anderen Versuche können entsprechend des Versuchsskriptes durchgeführt werden. Die Kursbetreuer nehmen dabei eine beratende Funktion ein. Im Versuch Zellkernfärbung wird der Makronukleus (vgl. Abschnitt 1.1.3) von *T. thermophila* mittels Methylgrün sichtbar und die Phagocytose kann innerhalb kurzer Zeit durch Fütterung von *T. thermophila* mit Tusche beobachtet werden. Aufgabe im Versuch zur Konjugation ist es, den Paarungstyp einer unbekannten Kultur von

T. thermophila durch Beobachtung seines Paarungsverhaltens mit den bekannten Paarungstypen I bis VII zu bestimmen. Zu diesem Zweck wird der den Schülern unbekannte Paarungstyp X bereits etwa drei Stunden vor Laborkursbeginn gemeinsam mit den anderen Paarungstypen in eine Multiwellplatte pipettiert (Abb. 33).

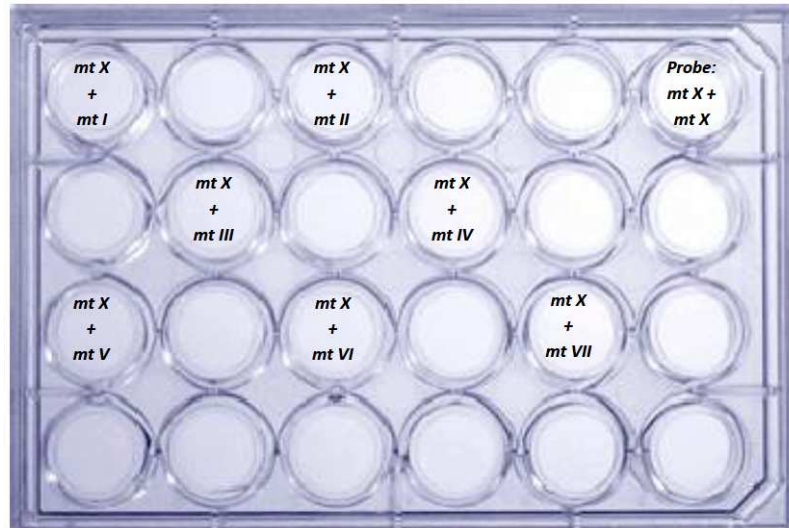


Abbildung 33: Pipettierschema der Paarungstypen von *T. thermophila* in der Multiwellplatte (mt=mating type, X=unbekannter Paarungstyp, Abbildung aus Kaminska 2010)

Nur in der *well*, in der der unbekannte Paarungstyp dem bekannten Paarungstyp entspricht, ist keine Konjugation zu beobachten. Um Wartezeiten für besonders schnelle Schüler zu vermeiden, wird optional noch die Zelltiterbestimmung einer *T. thermophila*-Kultur mittels THOMA-Zählkammer angeboten. Abschließend werden im Seminarraum des BeLL Bio die Beobachtungen der Schüler, sowie die während der Versuchsdurchführung aufgetretenen Schwierigkeiten, besprochen. Sollten einzelne Arbeitsgruppen bestimmte Phänomene nicht beobachtet haben, werden diese als vorbereitete Fotos im Rahmen einer PowerPoint-Präsentation veranschaulicht.

Tabellarische Planung des Laborkurses Einzeller

In der nachfolgenden Tabelle 67 wird ein Überblick über den Ablauf der Laborveranstaltung Einzeller gegeben.

Tabelle 67: Tabellarische Planung des Laborkurses Einzeller (BUW=Bergische Universität Wuppertal, LP=Lehrperson, PPP=PowerPoint-Präsentation, SuS=Schülerinnen und Schüler)

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
09:00h	Begrüßung (Haupteingang der BUW)	Schülergruppen werden vom Laborkursleiter am Haupteingang der BUW in Empfang genommen. Anschließend wird gemeinsam der Seminarraum des BeLL Bio aufgesucht.	Vorstellung des Laborkursleiters. Begrüßung an der Universität zur Betonung der authentischen Lernumgebung. Treffen von Absprachen mit der begleitenden LP bezüglich verspäteter SuS. Erste Einschätzung der Aufmerksamkeit und Motivation der SuS möglich.	---
09:15h	Einstieg (Seminarraum BeLL Bio)	Verständnisprobleme der vier im Laborkurs durchzuführenden Versuche werden erörtert.	Gelenktes Unterrichtsgespräch.	Versuchsskript
09:30h	Sicherheitsbelehrung + Einweisung in Mikroliterpipetten und Mikroskope (Seminarraum + Laborraum BeLL Bio)	Laborkursleiter führt eine verpflichtende Sicherheitsbelehrung durch, die von den SuS zu unterzeichnen ist. Laborkursleiter führt Mikroliterpipetten vor. Verteilen der Laborkittel, Einweisung in die Mikroskope und Belegung der Arbeitsplätze.	Gewährleistet sicheres Arbeiten im Labor und sicheren Umgang mit den im Labor zu verwendenden Arbeitsgeräten. Persönliche Verantwortung für einen Arbeitsplatz erhöht die Sorgfalt.	PPP, Beamer Mikroliterpipetten Lichtmikroskope
09:45h	Erarbeitung I (Laborraum BeLL Bio)	SuS führen Pipettierübung durch.	SuS eignen sich selbstständig in 2er-Gruppen eine wichtige Voraussetzung für die erfolgreiche Durchführung des Versuches an.	Mikroliterpipetten, Versuchsskript, Sicherheitsausrüstung (Kittel, Schutzbrillen, Handschuhe)
09:55h	Erarbeitung II (Laborraum BeLL Bio)	SuS führen die folgenden vier Versuche unter Einsatz der Lichtmikroskopie durch: <ul style="list-style-type: none">• Versuch: Smoking <i>Euglena</i>• Versuch zur Konjugation von <i>T. thermophila</i>• Versuch zur Zellkernfärbung bei <i>T. thermophila</i>	Arbeit in 2er-Gruppen. Reihenfolge der Versuche wird nicht vorgegeben. Offene Reihenfolge erfordert Absprache der Gruppen untereinander, da nicht sämtliche Materialien für alle Gruppen gleichzeitig verfügbar sind.	Lichtmikroskope Mikroliterpipetten Multiwellplatten Wasserstrahlpumpe

Zeit	Phase (Ort)	Unterrichtsaktivität	Didaktisch-methodischer Kommentar	Medien
		<ul style="list-style-type: none"> Versuch zur Phagocytose bei <i>T. thermophila</i> 		u.a.
12:10h	Ergebnis-sicherung (Seminarraum des BeLL Bio)	<p>Laborkursleiter erfragt Ergebnisse der SuS-Gruppen und zeigt parallel vorbereitete mikroskopische Aufnahmen der Versuche.</p> <p>Diskussion über die Ergebnisse und aufgetretene Probleme.</p>	<p>Gelenktes Unterrichtsgespräch.</p> <p>Laborkursleiter gibt Anregungen durch gezieltes Nachfragen.</p>	PPP, Beamer
12:40h-13:00h	Posttest (Seminarraum BeLL Bio)	SuS füllen Fragebogen aus.	Bestandteil des Pre-, Posttest-Designs	Fragebogen

2.2.5 Vorgehensweise bei der Dateneingabe und Auswertung

2.2.5.1 Verwendete Hardware und Software zur Dateneingabe

Die in der vorliegenden Arbeit eingesetzten Fragebögen wurden mit einem Fujitsu fi-5110C Datenscanner eingelesen und mit der Software FormPro 2.5. digitalisiert. Den für OCR (*Optical Character Recognition*)-Software üblichen Problemen bei der Erkennung von Handschrift und Korrekturen wurde durch visuelle Kontrolle jedes eingescannten Datensatzes begegnet.

Für die Videoaufzeichnungen vom Experimentierverhalten der Schüler im Laborkurs Genetischer Fingerabdruck wurden von Dahmen (2012) zwei Panasonic HDC TM 60 Kameras ohne zusätzliche externe Mikrofone verwendet. Die Videosequenzen wurden im .mts-Format aufgezeichnet.

2.2.5.2 Verwendete Software zur Datenauswertung

Für die Auswertung der Fragebögen wurde in erster Linie die Software IBM SPSS Statistics 20 eingesetzt. Zur Erstellung und Überprüfung eines Strukturgleichungsmodells wurde IBM AMOS 20 verwendet. Die Analyse der Fragebogendaten erfolgte in zwei Schritten. Zur Optimierung der Messqualität wurde zunächst eine Itemanalyse (Häußler et al. 1998, Bortz & Döring 2003) für die Wissensitems und die affektiven Items vorgenommen. In einem zweiten Schritt erfolgte die deskriptive und schließende Statistik.

2.2.5.3 Itemanalyse der Wissensitems

In den eingesetzten Fragebögen sollen die Wissensitems die Einschätzung der kognitiven Leistungsfähigkeit auf einem bestimmten Wissensgebiet ermöglichen. Für jedes richtig beantwortete Item wurde ein Punkt vergeben, für jedes falsch beantwortete Item kein Punkt. Generell spricht eine hohe erreichte Punktzahl im Wissenstest für eine hohe, eine niedrigere Punktezahl für eine geringere kognitive Leistungsfähigkeit. Problematisch ist ein zu leichter oder ein zu schwerer Wissenstest. Ein sehr leichter Test erlaubt es nicht, die kognitive Leistungsfähigkeit von Schülern einzuschätzen, da jeder, sogar die leistungsschwächeren Schüler, die zu leichten Testitems richtig beantworten können. Die sehr guten Schüler hingegen können ihr kognitives Potential gar nicht ausschöpfen, da sie eigentlich deutlich mehr leisten könnten. Ein zu schwerer Test bietet ebenfalls keine Möglichkeit, die kognitive Leistungsfähigkeit von Schülern einzuschätzen. Weder den leistungsschwächeren, noch den leistungsstärkeren Schülern gelingt es, einen solchen Test zu lösen. Ein zu schwerer Test vermittelt somit den Eindruck, dass die kognitive Leistungsfähigkeit der teilnehmenden Schüler geringer ist als dies tatsächlich der Fall ist. Es gilt also, zu leichte und zu schwere Items zu identifizieren und aus der weiteren Auswertung zu entfernen (Häußler et al. 1998). Dies geschieht durch Berücksichtigung des Schwierigkeitsindex und des Trennschärfekoeffizienten.

Als Maß für die Schwierigkeit eines Items dient der Schwierigkeitsindex. Dieser gibt die prozentuale Lösungshäufigkeit eines Testitems bezogen auf alle Probanden an (Zöfel 2002). In Orientierung an vergleichbaren Studien (Leibold 1997, Wilde 2004, Großschedl & Harms 2008) fließen in die vorliegende Auswertung ausschließlich Items ein, die im Nachtest einen Schwierigkeitsindex von 10 % bis 90 % aufweisen. Tabelle 68 gibt einen Überblick über die ursprüngliche sowie die verbleibende Itemzahl nach Bestimmung des Schwierigkeitsindex für die kognitiven Messinstrumente der drei evaluierten Laborkurse.

Tabelle 68: Verbleibende Itemzahl in den Wissenstests der drei Laborkurse nach Bestimmung des Schwierigkeitsindex

Wissenstest	Ursprüngliche Itemzahl	Itemzahl nach Bestimmung des Schwierigkeitsindex
Genetischer Fingerabdruck	106	73
LDH	102	73
Einzeller	109	85

Als wichtigstes Kriterium zur Beurteilung der Brauchbarkeit eines Items gilt neben dem Schwierigkeitsindex der Trennschärfekoeffizient. Dieser wird nur für die Items bestimmt, die nach Bestimmung des Schwierigkeitsindex noch in der Auswertung verblieben sind.

Der Trennschärfekoeffizient gibt für jedes Item an, wie gut es ihm gelingt, zwischen „guten“ und „schlechten“ Testteilnehmern zu trennen. Er entspricht der Korrelation zwischen der Aufgabenantwort und der im Test erreichten Gesamtpunktzahl. Der Trennschärfekoeffizient ist bei mittleren Schwierigkeitsindices am höchsten, während er bei hohen und niedrigen Schwierigkeitsindices abfällt (Zöfel 2002). Aufgaben mit einer Trennschärfe <0.1 wurden zur Reliabilitätsoptimierung der Items aufgespürt und aus der weiteren Auswertung entfernt (vgl. Häußler et al. 1998). Tabelle 69 gibt an, wie viele Items der drei Wissenstests nach diesem Schritt noch in der Auswertung verbleiben. Ebenfalls finden sich hier die Reliabilitäten dieser Tests.

Tabelle 69: Verbleibende Itemzahl in den Wissenstests der drei Laborkurse nach Bestimmung des Schwierigkeitsindex und der Trennschärfe einschließlich der Reliabilität (Cronbachs α)

Wissenstest	Itemzahl nach Bestimmung des Schwierigkeitsindex	Itemzahl nach Bestimmung der Trennschärfe	Cronbachs α
Genetischer Fingerabdruck	73	56	0,866
LDH	73	50	0,829
Einzeller	85	54	0,856

2.2.5.4 Itemanalyse der affektiven Messinstrumente

Der Ausschluss von Items geringer Trennschärfe wurde auch für die affektiven Messinstrumente vorgenommen. Bei diesen gibt die Trennschärfe nicht – wie bei den Wissensitems – an, wie gut ein Item zwischen guten und schlechten Schülern differenziert. Vielmehr lässt die Trennschärfe bei den affektiven Items erkennen, wie gut ein Item zwischen Personen mit hoher und Personen mit niedriger Merkmalsausprägung trennt. Affektive Items mit einer Trennschärfe von $<0,26$ wurden aus der Auswertung ausgeschlossen. Die entsprechenden Items sind bei der Beschreibung der affektiven Messinstrumente unter Punkt 2.2.3.2.3 grau hinterlegt. Auch hier wurden nur die in der Auswertung verbliebenen Items für die Berechnung der berichteten Reliabilitäten verwendet.

2.2.5.5 Prüfung der Daten auf Normalverteilung

Die Auswahl geeigneter Methoden deskriptiver und schließender Statistik setzt für intervallskalierte Variablen die Berücksichtigung der Datenlage, insbesondere die Ermittlung der Rohdatenverteilung, voraus. Während Intervallskalierung für die Punkte in einem Wissenstest angenommen werden kann, ist diese für die Daten, welche mittels Likert-Skala erhoben wurden, strittig. In Anlehnung an Hadler (2005) werden in der vorliegenden Studie auch die in den affektiven Messinstrumenten eingesetzten fünfstufigen Likert-Skalen als intervallskaliert betrachtet. Häufig werden normalverteilte Daten

angestrebt, da diese von vielen inferenzstatistischen Verfahren vorausgesetzt werden (Bortz & Döring 2003).

In der vorliegenden Studie werden die Schiefe (Skewness) und die Wölbung (Kurtosis) zur Beurteilung der Daten auf Normalverteilung herangezogen. Die Skewness beschreibt wie weit die Häufigkeitsverteilung von einer symmetrischen Verteilung abweicht. Im Falle einer Normalverteilung ist die beobachtete Schiefe null (Bühl 2010). Liegt das Maximum der unimodalen Verteilung weiter auf der linken Seite, so spricht man von einer rechtsschiefen Verteilung. Die berechnete Schiefe ist in diesem Fall positiv. Bei einer linksschiefen Verteilung liegt das Maximum weiter rechtsseitig; die Schiefe ist negativ. Über- beziehungsweise unterschreitet die Schiefe die Grenzwerte von +3 beziehungsweise -3, so ist die Verteilung als nicht mehr hinreichend normalverteilt für die Anwendung parametrischer Verfahren anzusehen (vgl. Kline 2011).

Neben der Schiefe spielt die Kurtosis (Exzess/Wölbung) für die Einschätzung der Normalverteilung eine Rolle. Diese ist im Falle einer breitgipfeligen (hoher Wert) oder schmalgipfeligen Verteilung gegeben (Bühl 2010). Bei normalverteilten Daten liegt definitionsgemäß eine Kurtosis von 3 vor. Diese Version der Kurtosisangabe wird auch als Pearson-Kurtosis bezeichnet. In vielen Statistikprogrammen – so auch in SPSS und AMOS – wird grundsätzlich der Wert 3 von der Wölbungsstatistik subtrahiert (Curran et al. 1996, Baltes-Götz 2010). Diese auf 0 zentrierte Kurtosis wird auch als Fisher-Kurtosis⁶ bezeichnet. Überschreitet die Fisher-Kurtosis den Wert 7 (vgl. Kline 2011), so liegt eine gravierende Abweichung von der Normalverteilung vor.

Schiefe und Wölbung der erhobenen Daten liegen innerhalb der beschriebenen Grenzwerte. Für die weitere Auswertung werden daher parametrische Verfahren eingesetzt. Auch durch die in einigen Vergleichsgruppen etwas höheren Abweichungen in den Skalen Transparenz und Druck ist keine Überbetonung nicht vorhandener Unterschiede durch die Anwendung parametrischer Tests zu befürchten. Letztendlich ist davon auszugehen, dass bei Verletzung der Voraussetzung der Normalverteilung ein parametrisches Verfahren an Teststärke verliert. Eine bestätigte Hypothese würde daher erst recht bestätigt werden, wenn die Voraussetzung für die Durchführung eines parametrischen Verfahrens ideal wären (Bortz et al. 2008).

Von der Verteilung der Daten hängt ab, welche statistischen Kennwerte (z.B. Mittelwert vs. Median) zu ihrer Beschreibung verwendet werden können (Zöfel 2002). Wie im vorhergehenden Abschnitt festgestellt wurde, sind die Daten ausreichend normalverteilt, so dass hier Mittelwerte und Standardabweichungen zur Beschreibung der Ausprägung einer Variablen verwendet werden

⁶ Sämtliche Ausführungen bzgl. der Kurtosis beziehen sich in der vorliegenden Studie auf die Fisher-Kurtosis.

können. Um die Stärke des Zusammenhanges zwischen zwei Variablen festzustellen, werden Korrelationen verwendet.

2.2.5.6 Prüfung der Daten auf Korrelationen

Die Stärke eines linearen Zusammenhangs zwischen zwei Variablen kann über den Korrelationskoeffizienten r beschrieben werden. Dieser liegt zwischen -1 und $+1$. Bei einem perfekten gleichsinnigen Zusammenhang (hohe Werte einer Variable korrelieren mit hohen Werten der anderen Variable) ist $r=+1$. Im Falle eines perfekten negativen Zusammenhangs ist $r=-1$. Liegt der Korrelationskoeffizient bei 0 , so besteht kein (linearer) Zusammenhang zwischen den Variablen (Zöfel 2002, Abb. 34).

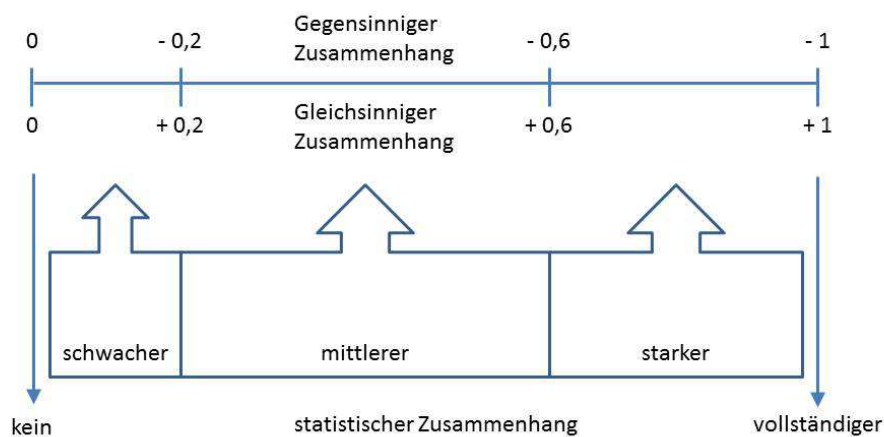


Abbildung 34: Die Interpretation des Korrelationskoeffizienten (verändert nach Quatember 2011)

Ist $|r| < 0,2$ so spricht Quatember (2011) von kleinen statistischen Zusammenhängen. Liegt $|r|$ zwischen $0,2$ und $0,6$ spricht er von mäßigen und ab $r > 0,6$ von starken Korrelationen. Bühl (2010) spricht dagegen erst ab $|r| > 0,7$ von hohen und ab $|r| > 0,9$ von sehr hohen Korrelationen. Die Berechnung des Korrelationskoeffizienten hängt vom Skalenniveau der betreffenden Variablen ab. Für intervall- und normalskalierte Variablen wird in der vorliegenden Studie der Pearson-Korrelationskoeffizient verwendet. Für nicht normalverteilte oder ordinalskalierte Variablen wird auf Kendalls Tau zurückgegriffen (Bühl 2010). Kendalls Tau findet konkret unter Punkt 3.2.1.5 bei der Korrelierung der nicht normalverteilten Videodaten mit den Fragebögen sowie unter Punkt 3.2.2.1.1 bei der Korrelierung des dispositionalen Interesses mit ordinalskalierten Schulnoten Verwendung.

2.2.5.7 Statistische Tests – AN(C)OVA, Post-Hoc-Tests

In der vorliegenden Studie wurden sowohl Intergruppenvergleiche zwischen verschiedenen Gruppen innerhalb eines Messzeitpunktes vorgenommen als auch Intra- und Intergruppenvergleiche über

mehrere Messzeitpunkte. Statistische Tests innerhalb eines Messzeitpunktes werden nachfolgend auch als Querschnittsvergleiche, Tests über mehrere Zeitpunkte als Längsschnittanalysen bezeichnet.

Für Querschnittsvergleiche zwischen zwei Mittelwerten wurde der t-Test für unabhängige Stichproben angewendet. Tests auf bedeutsame Unterschiede zwischen mehr als zwei Mittelwerten wurden mittels einfaktorieller ANOVA nach dem Allgemeinen Linearen Modell (ALM) in SPSS durchgeführt.

Einige Variablen wie das Fachwissen, das biologische Fähigkeitsselbstkonzept, die experimentbezogene Selbstwirksamkeitserwartung oder das akute Interesse wurden zu mehreren Messzeitpunkten erhoben. Folgendes Vorgehen wurde für die Längsschnittvergleiche zwischen zwei oder mehreren Vergleichsgruppen (z.B. Versuchs- und Kontrollgruppe; männliche und weibliche Schüler) gewählt: Zunächst wurde mittels einfaktorieller ANOVA mit Messwiederholung kontrolliert, ob in der Kontrollgruppe eine zeitliche Veränderung über die drei Messzeitpunkte zu beobachten ist. Diese Überprüfung diente zum Ausschließen von Pretesteffekten. Bei diesen handelt es sich um zeitliche Veränderungen der abhängigen Variable, die nicht durch das eigentliche Treatment (z.B. den Laborbesuch), sondern allein durch das Ausfüllen des Pretests bedingt werden. Anschließend wurde zunächst ein t-Test für unabhängige Stichproben (bei zwei Vergleichsgruppen) oder eine einfaktorielle univariate Varianzanalyse (bei mehr als zwei Vergleichsgruppen) zwischen den Vergleichsgruppen zum Pretestzeitpunkt durchgeführt, um bereits vor dem Laborbesuch bestehende Unterschiede aufzudecken. Im Falle von Pretestäquivalenz wurden zweifaktorielle ANOVAs mit Messwiederholung berechnet. Der erste Faktor der ANOVA ist dabei die zeitliche Veränderung und der zweite Faktor derjenige, der die Vergleichsgruppe definiert (z.B. Treatment, Geschlecht, Migrationshintergrund). Bei der Ergebnisinterpretation einer zwei- oder mehrfaktoriellen ANOVA ist zwischen Haupt- und Interaktionseffekten zu unterscheiden (Rasch et al. 2010b). Mit Haupteffekt ist der Einfluss eines Faktors (z.B. Zeit) auf eine abhängige Variable (z.B. das Wissen) gemeint. Ein signifikanter Haupteffekt Zeit sagt lediglich aus, ob – unabhängig von der Vergleichsgruppe – irgendeine zeitliche Veränderung über die Messzeitpunkte beobachtet werden kann. Dieser Haupteffekt kann auch signifikant werden, wenn sich zwei Vergleichsgruppen gleichermaßen über die Zeit verändern. Der Effekt allein sagt somit noch nichts darüber aus, ob sich zwei Vergleichsgruppen zeitlich unterschiedlich entwickeln und ob die Gruppenzugehörigkeit einen Einfluss auf die Entwicklung hat. Der Haupteffekt Treatment (damit können in der vorliegenden Studie Vergleiche von Versuchs- und Kontrollgruppe, aber auch Subgruppenvergleiche nach Geschlecht und Migrationshintergrund gemeint sein) sagt aus, ob die Vergleichsgruppen sich bedeutsam unterscheiden. Ein signifikanter Haupteffekt Treatment kann aber bei einer Messwiederholung beispielsweise auch bedeuten, dass ein Unterschied zwischen den

Vergleichgruppen bereits zum Pretestzeitpunkt besteht und über mehrere Messzeitpunkte erhalten bleibt. Dieser Haupteffekt allein sagt somit wenig darüber aus, ob sich zwei Gruppen zeitlich unterschiedlich entwickeln. Mit dem Interaktionseffekt, auch Wechselwirkung der beiden Faktoren genannt, ist in dieser Arbeit den gemeinsamen Einfluss der zwei Faktoren Zeit und Treatment auf die abhängige Variable gemeint. Er gibt also im vorliegenden Fall Auskunft darüber, ob sich die untersuchten Gruppen über die Zeit signifikant unterschiedlich entwickeln. Somit ist der Interaktionseffekt für die Beurteilung des Einflusses des Laborbesuches von besonderem Interesse. Um die Ergebnisse der Längsschnittanalysen besser nachvollziehbar zu machen, wird bei zweifaktoriellen ANOVAS grundsätzlich angegeben, ob ein Haupt- oder Interaktionseffekt berichtet wird. Bei Vorliegen eines signifikanten Interaktionseffektes wurden in der Regel paarweise ANOVAs (Pre-/Posttest & Pre-/Follow-up-Test) durchgeführt, um festzustellen, zwischen welchen Messzeitpunkten die beobachtete Unterschiede liegen.

Lagen bereits zum Zeitpunkt des Pretest Unterschiede zwischen den Vergleichsgruppen vor, so erschien oben genanntes Vorgehen nicht adäquat, da Pretestunterschiede Veränderungsraten der abhängigen Variablen beeinflussen können (Kaluza & Schulze 2000). In diesem Fall wurden Kovarianzanalysen (ANCOVAs) mit der unabhängigen Vergleichsvariable (z.B. Geschlecht) als Faktor, den Pretestergebnissen (z.B. dem Vorwissen) als Kovariate und dem Post- beziehungsweise den Follow-up-Test-Resultaten als abhängige Variable vorgenommen. Im Rahmen der ANCOVA wird der Einfluss der Kovariate aus der abhängigen Variablen herauspartialisiert. Da für den Laborkurs Einzeller kein Follow-up-Test durchgeführt wurde, beschränken sich entsprechende Analysen auf die ersten beiden Messzeitpunkte. Neben den Vergleichen von Versuchs- und Kontrollgruppe wurden weitere Subgruppenanalysen auf Unterschiede in Abhängigkeit von Geschlecht, Migrationshintergrund und beruflicher Orientierung durchgeführt. Diese erfolgten in Anlehnung an der zuvor beschriebenen Auswerteprozedur.

Ein signifikantes Ergebnis in einer Varianzanalyse besagt lediglich, dass zwischen zwei Gruppen ein signifikanter Unterschied besteht, es sagt aber nichts darüber aus, zwischen welchen Gruppen dieser vorliegt. Aus diesem Grunde wurden bei Vergleichen von mehr als drei Gruppen im Falle eines signifikanten Ergebnisses in der ANOVA oder ANCOVA Post-Hoc-Tests angeschlossen. Diese testen alle Vergleichsgruppen gegeneinander. Im Falle der ANOVA wurden Post-Hoc-Tests nach Scheffé, im Falle der ANCOVA nach Bonferroni durchgeführt. Generell wäre auch die Durchführung vieler vergleichender t-Tests statt der ANOVAs denkbar, doch dieses Vorgehen ist mit dem Problem der α -Fehlerkumulierung behaftet. Der α -Fehler, also die Wahrscheinlichkeit, die Nullhypothese abzulehnen, obwohl sie wahr ist, wird deutlich größer mit der Anzahl der durchgeführten statistischen Tests. Der α -Fehler ergibt sich aus

$$\alpha_{\text{gesamt}} = 1 - (1 - \alpha_{\text{Test}})^m \quad (\text{Rasch et al. 2010b})$$

Dabei ist α_{gesamt} das kumulierte α -Niveau, α_{Test} das α -Niveau (=Irrtumswahrscheinlichkeit) des einzelnen Tests und m die Anzahl der durchgeführten Einzeltests. Für den Vergleich von drei Mittelwerten wären drei t-Tests notwendig. Zwar wird in der Regel der α -Fehler für jeden Test auf 5 % ($p=0,05$) festgelegt, die Gesamtfehlerwahrscheinlichkeit liegt nach oben genannter Formel aber bereits bei 14 %. Entsprechend erscheint es wenig sinnvoll, Vergleiche von zwei oder mehr Gruppen über mehrere Messzeitpunkte mit zahlreichen t-Tests vorzunehmen (Rasch et al. 2010b).

2.2.5.7.1 Festlegen der Signifikanzniveaus

Statistische Tests geben die Irrtumswahrscheinlichkeit p an. Dies ist die Wahrscheinlichkeit, aufgrund seiner Stichprobendaten einen bedeutsamen Unterschied anzunehmen, der in der Realität nicht gegeben ist. Klassischerweise wird ab einer Irrtumswahrscheinlichkeit $p \leq 0,05$ von einem signifikanten Ergebnis gesprochen. Die im Ergebnisteil verwendete Abstufung der Signifikanzniveaus orientiert sich an Zöfel (2002, Tab. 70).

Tabelle 70: Signifikanzniveaus und deren Kennzeichnung in Abbildungen (vgl. Zöfel 2002)

Signifikanzniveau	Bezeichnung	Kennzeichnung in Abbildungen
$p \leq 0,05$	signifikant	*
$p \leq 0,01$	sehr signifikant	**
$p \leq 0,001$	höchst signifikant	***

In Anlehnung an vergleichbare Untersuchungen wird in der vorliegenden Studie ein Signifikanzniveau von $0,05 < p \leq 0,1$ als tendenzieller Unterschied bezeichnet (vgl. Mittag & Jerusalem 1999, Hampel et al. 2007). Entsprechend der Darstellung in Tabelle 70 werden signifikante Unterschiede in den Grafiken des Ergebnisteils durch „*“ gekennzeichnet. Diese Kennzeichnung wird nicht für Längsschnittanalysen verwendet, da Interaktionen sich so kaum sinnvoll in einer Abbildung hervorheben lassen.

2.2.5.7.2 Untersuchung der Effektstärke

Während die Signifikanz eines Tests nur etwas über das Vorliegen oder Nicht-Vorliegen eines Unterschiedes bei einer gewissen Irrtumswahrscheinlichkeit aussagt, gibt die Effektgröße Auskunft darüber, wie stark der Einfluss unabhängiger Variablen auf die abhängige(n) Variable(n) ist. Zudem ist die Effektstärke – im Gegensatz zur Signifikanz – unabhängig von der Stichprobengröße. Als Effektgröße finden in der vorliegenden Studie für unabhängige t-Tests die Effektgrößen Omega-

Quadrat (ω^2 , Albert & Koster 2002) und für abhängige t-Tests sowie Varianzanalysen das partielle Eta-Quadrat (η^2) Verwendung. Die Berechnung der Effektgröße ω^2 erfolgt aus dem von SPSS- ausgegebenen t-Wert und den verglichenen Stichprobengrößen N_1 und N_2 nach:

$$\omega^2 = (t^2 - 1) / (t^2 + N_1 + N_2 - 1) \quad (\text{Albert \& Koster 2002})$$

Für die Beurteilung von ω^2 werden folgende Richtwerte berücksichtigt (Tabelle 71)

Tabelle 71: Richtwerte zur Beurteilung von ω^2 (nach Albert & Koster 2002)

Wert	Interpretation
$\omega^2 \leq 0$	kein Effekt
$0,01 \leq \omega^2 < 0,06$	geringer Effekt
$0,06 \leq \omega^2 < 0,15$	mittlerer Effekt
$\omega^2 \geq 0,15$	erheblicher Effekt

Das partielle η^2 wird für abhängige t-Tests berechnet nach:

$$\eta^2 = \frac{(t_{\text{abhängig}}^2 / df)}{1 + (t_{\text{abhängig}}^2 / df)} \quad (\text{Rasch et al. 2010a})$$

Für Varianzanalysen wird das partielle η^2 entsprechend der SPSS-Ausgabe verwendet. Die Beurteilung der Effektgröße orientiert sich an den Angaben von Cohen (1988, Tabelle 72):

Tabelle 72: Richtwerte zur Beurteilung von η^2 (Cohen 1988)

Wert	Interpretation
$\eta^2=0,01$	Kleiner Effekt
$\eta^2=0,06$	Mittlerer Effekt
$\eta^2=0,14$	Großer Effekt

2.2.5.8 Bildung von Subgruppen durch hierarchische Clusteranalyse

Eine hierarchische Clusteranalyse findet in der vorliegenden Hauptstudie Anwendung, um die Teilnehmer der BeLL Bio-Kurse und der Kontrollgruppe anhand ähnlicher Ausprägungen ihrer affektiven Personenvariablen in Lernertypen einzuteilen. Berücksichtigt wurden für die Analyse explizit die im Pretest erhobenen Variablen Fachinteresse, Sachinteresse, Verständnis,

Wissenschaftsverständnis, FSWEx.V1-Planung, FSWEx.V1-Durchführung, FSWEx.V1-Auswertung und biologisches Fähigkeitsselbstkonzept.

Bei der hierarchischen Clusterbildung wird zunächst jeder Fall (Schüler) als einzelner Cluster behandelt und die Distanz zwischen den Clustern bestimmt. Als gängiges Distanzmaß wird hier der *quadrierte Euklidische Abstand*, als Cluster-Methode die *Complete Linkage* (Entferntester Nachbar) verwendet. Bei diesem Clusterverfahren ist die Distanz zwischen zwei Clustern diejenige zwischen dem entferntesten Fallpaar. Bei der Paarbildung stammt jeweils ein Fall aus einem der beiden Cluster (Bühl 2010). Die zwei Cluster mit der geringsten Distanz werden somit in einem größeren gemeinsamen Cluster vereinigt. Diese Prozedur wird so oft wiederholt, bis globalere Cluster mit großer Distanz zueinander identifiziert werden. Diese werden verwendet, um die Schüler in Gruppen einzuteilen. Im Rahmen der Analyse konnten drei Cluster unterschiedlicher persönlicher Dispositionen identifiziert werden, die im Ergebnisteil dieser Arbeit als Schülertypen bezeichnet werden.

2.2.5.9 Faktorenanalytische Prüfung der Dimensionalität des Fragebogens zur experimentbezogenen Selbstwirksamkeitserwartung

Mittels der Faktorenanalyse ist es möglich, eine größere Anzahl an Variablen auf eine kleinere Anzahl hypothetischer, latenter, nicht beobachtbarer Größen – Faktoren genannt – zurückzuführen (Zöfel 2002). Für die erst im folgenden Abschnitt vorgestellte Strukturgleichungsmodellierung ist es bereits an dieser Stelle sinnvoll, zwischen den zwei Grundtypen der Faktorenanalyse zu differenzieren: der Explorativen Faktorenanalyse (EFA) und der Confirmatorischen Faktorenanalyse (CFA). Die EFA wird eingesetzt, wenn die Beziehung zwischen den beobachtbaren und den latenten Variablen unbekannt oder unsicher ist. Sie erlaubt es, die minimale Zahl der hinter den beobachteten Variablen stehenden latenten Variablen zu identifizieren. Dabei werden die beobachtbaren Variablen, die gemeinsam eine latente Variable abbilden, einem Faktor zugeordnet. Häufig wird die EFA bei der Entwicklung eines neuen Messinstrumentes eingesetzt. Sie ermöglicht es hier zu prüfen, ob die Items, welche formuliert wurden um eine bestimmte Variable erfassen, auch tatsächlich eine besonders hohe Faktorladung auf dieser latenten Variable aufweisen (Byrne 2010).

In der vorliegenden Studie kommt eine EFA im Rahmen der Entwicklung des Messinstrumentes FSWEx zum Einsatz um zu prüfen, ob die Items des FSWEx tatsächlich wie beabsichtigt drei unterschiedliche Messdimensionen (Planung, Durchführung, Auswertung) abbilden oder nur eine einzige. Vor Durchführung der Faktorenanalyse ist die Eignung der zu berücksichtigenden Variablen zu überprüfen. Dies geschah unter Anwendung des Barlett-Tests. Dieser überprüft, ob es sein könnte,

dass in der Grundgesamtheit – trotz der vorliegenden Korrelationskoeffizienten in der Stichprobe – keine bedeutsamen Korrelationen vorliegen (Bühl 2010). Ein signifikantes Testergebnis bedeutet, dass die untersuchten Variablen mehr als zufällig korrelieren und diese daher für eine Faktorenanalyse geeignet erscheinen. Als weiteres Maß für die Güte des Faktorenmodells kann das Kaiser-Meyer-Olkin (KMO)-Maß verwendet werden. Dieses kann Werte zwischen 0 und 1 annehmen. Ein KMO unter 0,6 spricht gegen die Verwendung der Variablen für eine Faktorenanalyse (Bühl 2010). Um zu prüfen, ob die Items einer FSWE-Subskala tatsächlich dasselbe Konstrukt erfassen, wurde nach Prüfung der eben beschriebenen Voraussetzungen eine explorative Faktorenanalyse als Hauptkomponentenanalyse mit Varimax-Rotation durchgeführt. Idealerweise sollten die FSWE-Items, welche eine gemeinsame Teilkompetenz abbilden sollen, die höchsten Faktorladungen auf einem gemeinsamen Faktor aufweisen. Die Resultate dieser Analysen finden sich unter Punkt 3.2.1 im Ergebnisteil.

Im Gegensatz zur EFA kommt die CFA zum Einsatz, wenn die Beziehung – sei es gestützt durch Theorie oder empirische Untersuchungen – zwischen den beobachtbaren und latenten Variablen bereits bekannt ist (Byrne 2010). Sie testet also ein theoretisch angenommenes Modell (Bühl 2010) und ist damit der für die Strukturgleichungsmodellierung relevante Grundtyp der Faktorenanalyse. Die Bedeutung der Strukturgleichungsmodellierung für die vorliegende Studie wird im folgenden Abschnitt dargestellt.

2.2.5.10 Erstellung eines Modells der affektiven Einflussgrößen am Lernort BeLL Bio

Die bisher vorgestellten Verfahren erlauben es nicht, die komplexe Beziehungsstruktur der lernwirksamen Einflussfaktoren am Lernort BeLL Bio hinreichend zu beschreiben. Insbesondere die Überprüfung des unter Punkt 1.2.5 formulierten Hypothesennetzes ist mit den bisher beschriebenen Methoden nicht ganzheitlich möglich. Weiterreichende Möglichkeiten bietet die lineare Strukturgleichungsmodellierung (Structural equation modeling, SEM), wie sie in der vorliegenden Studie in AMOS 20 umgesetzt wurde. Zum Verständnis des im Ergebnisteil vorgestellten linearen Strukturgleichungsmodells werden im Anschluss zunächst einige Grundbegriffe der Strukturgleichungsmodellierung vorgestellt.

In erster Näherung kann ein Strukturgleichungsmodell als Kombination von Regressions- und Faktorenanalyse aufgefasst werden. Gegenüber der linearen Regressionsanalyse liegt im SEM allerdings keine Beschränkung auf rekursive Modelle vor (Baltes-Götz 2010). Rekursive Modelle beinhalten nur einseitig gerichtete Kausalitätsbeziehungen und keine Feedback-Prozesse (Ammermann et al. 1975). Aufgrund der Möglichkeit, auch Wechselwirkungen zu erfassen, erscheint

das Strukturgleichungsmodell geeigneter die Realität abzubilden (Weiber & Mülhhaus 2010). Weiterhin bietet die Strukturgleichungsmodellierung – im Gegensatz zur linearen Regressionsanalyse – die Möglichkeit, latente Variablen abzubilden (Baltes-Götz 2010). Latente Variablen, auch Faktoren genannt, sind theoretische Konstrukte, welche sich nicht direkt beobachten und messen lassen. Hierzu zählen beispielsweise das Selbstkonzept und die Selbstwirksamkeit einer Person. Um diese Konstrukte dennoch erfassen zu können, bedient man sich beobachtbarer (=manifeste) Variablen, welche als messbare Indikatoren des Konstruktes dienen. Diese Indikatoren können zum Beispiel Punkte in einem Leistungstest oder affektive Einschätzungen auf einer Likert-Skala sein. Im Umgang mit Strukturgleichungsmodellen ist es hilfreich, zwischen endogenen und exogenen latenten Variablen zu unterscheiden. Exogene latente Variablen sind unabhängige Variablen, die – direkt oder indirekt – eine Veränderung endogener (abhängiger) latenter Variablen bewirken. Im SEM werden nur die Veränderungen der endogenen, nicht aber die Änderungen in den exogenen latenten Variablen erklärt (Byrne 2010).

Das traditionelle statistische Vorgehen, um den Zusammenhang eines Sets beobachtbarer und latenter Variablen zu untersuchen, ist die bereits im vorherigen Abschnitt vorgestellte Faktorenanalyse (Byrne 2010). Da die confirmatorische Faktorenanalyse (CFA) im Gegensatz zur explorativen Faktorenanalyse (EFA) ein theoretisch angenommenes Modell testet, ist sie das für die Strukturgleichungsmodellierung relevante Verfahren. Weil die CFA die Beziehung zwischen den latenten Faktoren und den manifesten, tatsächlich gemessenen Variablen beschreibt, wird sie im Strukturgleichungsmodell auch als reflektives Messmodell bezeichnet. (Byrne 2010). Das Strukturmodell bildet hingegen die vermuteten Zusammenhänge zwischen den latenten Variablen ab. Grafisch darstellen lässt sich das erstellte Kausalsystem in Form eines Pfaddiagramms (Weiber & Mülhhaus 2010, Abb. 35).

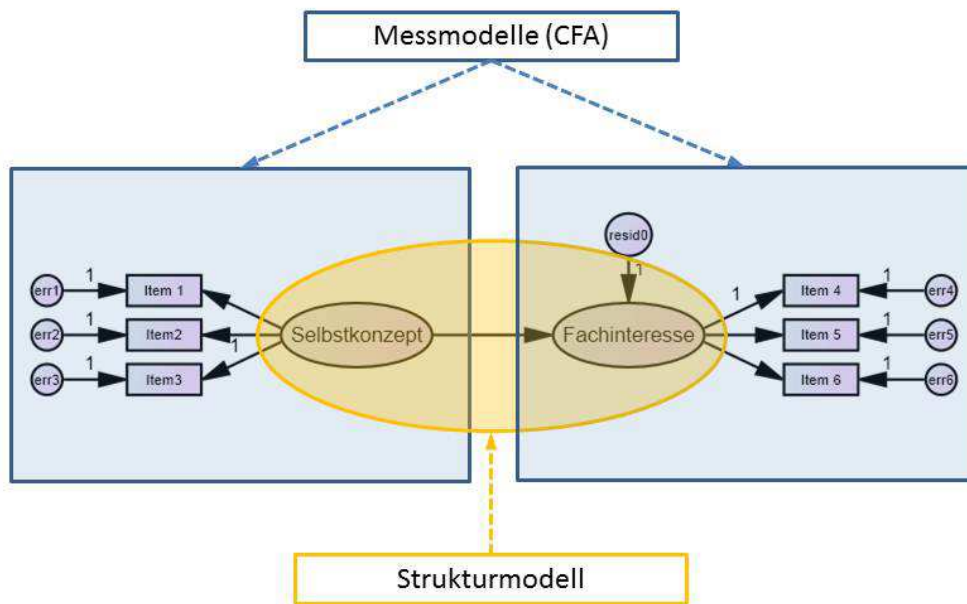







Abbildung 35: Pfaddiagramm für ein Strukturgleichungsmodell: Messmodelle (Confirmatorische Faktorenanalysen, CFA) und Strukturmodell sind gekennzeichnet (Abbildung in Anlehnung an Byrne 2010)

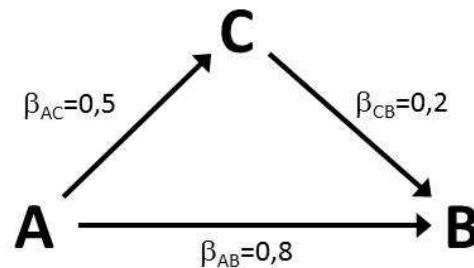
In dem im Ergebnisteil vorgestellten Pfaddiagramm (vgl. Punkt 3.2.2.5) werden latente Variablen als Ellipsen, manifeste Variablen als Rechtecke und Kausalbeziehungen als Pfeile dargestellt. Von exogenen Variablen gehen grundsätzlich nur Pfeile aus, während zu endogenen Variablen mindestens eine Pfeilspitze hinzeigt (Weiber & Mülhauß 2010). Doppelpfeile stellen Korrelationen dar und Kreise symbolisieren Fehlervariablen (Tab. 73).

Tabelle 73: Bedeutung der Symbole im Pfaddiagramm

Bedeutung	Grafische Darstellung im Pfaddiagramm
Latente Variable	
Manifeste Messvariable	
Residualvariable	
Kausalbeziehung	
Korrelationen	

Die verursachenden Variablen werden auch als Prädiktorvariablen, die durch sie beeinflussten endogenen Variablen auch als Kriteriumsvariablen bezeichnet. Die Stärke des Zusammenhangs zwischen zwei Variablen wird in Form der standardisierten Regressionsgewichte (β) angegeben. Ihr

Wertebereich liegt zwischen -1 und +1 (Weiber & Mülhhaus 2010). Standardisierte β -Gewichte mit einem Betrag $>0,2$ sind als bedeutungsvoll anzusehen (Chin 1998). Die Einschätzung des Einflusses einer Variable auf eine andere macht es erforderlich, zwischen direkten, indirekten und totalen Effekten zu differenzieren (Abb. 36).



Direkter Effekt von A auf B = $\beta_{AB}=0,8$

Indirekter Effekt von A auf B = $\beta_{AC} * \beta_{CB}=0,5*0,2=0,1$

Totaler Effekt von A auf B = Direkter Effekt + Indirekter Effekt = 0,9

Abbildung 36: Die Bedeutung direkter, indirekter und totaler Effekte im Strukturgleichungsmodell

Im oben angeführten Beispiel entspricht der direkte Effekt von A auf B dem standardisierten β_{AB} -Gewicht. Der indirekte Effekt von A auf B entspricht dem Produkt der β -Gewichte β_{AC} und β_{CB} . Der totale Effekt von A auf B ist die Summe aus direktem und indirektem Effekt (Kline 2011). Um den Einfluss der Veränderung einer Variablen auf eine andere abschätzen zu können, sind daher auch die indirekten Effekte von Bedeutung. Liegen mehrere indirekte Effekte vor, so werden diese addiert.

Der kovarianzanalytische Ansatz in AMOS versucht die Modellparameter der im Pfaddiagramm postulierten Beziehungen so zu schätzen, dass die auf den erhobenen Daten basierende Varianz-Kovarianz-Matrix möglichst genau vom Modell reproduziert werden kann (Weiber & Mülhhaus 2010). Als Schätzalgorithmus wurde in der vorliegenden Studie die Maximum Likelihood-Methode (ML) gewählt. Für die Parameterschätzung sieht Bagozzi (1981) eine Stichprobengröße von $N-t > 50$ als hinreichend an, wobei N die Stichprobengröße und t die Zahl der zu schätzenden Parameter ist. t beträgt im unter Punkt 3.2.2.5 vorgestellten Modell 158. Um eine ausreichend große Stichprobe für eine zuverlässige Parameterschätzung zu erhalten, wurden die Daten der Treatmentgruppen Genetischer Fingerabdruck und LDH gemeinsam in einem Modell ausgewertet ($N=301$). Dieses Vorgehen orientiert sich an vergleichbaren Studien (Pawek 2009) und erscheint zulässig, da sich die Wahrnehmung der beiden Laborkurse durch die Schüler in vielen Aspekten gleicht (vgl. Punkt 3.2.2.2.1). Zur Beurteilung der Modellgüte werden die in Tabelle 74 aufgeführten Kriterien verwendet.

Tabelle 74: Kriterien zur Beurteilung von Strukturgleichungsmodellen

Gütekriterium	Interpretation	Literaturquelle
Root-Mean-Square-Error of Approximation (RMSEA)	RMSEA ≤ 0,08 akzeptabler Modell-Fit	Browne & Cudeck (1993)
$\chi^2/\text{d.f.}$	$\chi^2/\text{d.f.} \leq 3$ akzeptabler Modell-Fit	Homburg & Giering (1996)
Incremental Fit Index (IFI)	IFI ≥ 0,90 guter Modell-Fit	Weiber & Mülhauß (2010)
Comparative Fit Index (CFI)	CFI ≥ 0,90 guter Modell-Fit	Homburg & Baumgartner (1995)

2.2.5.11 Videoanalytische Auswertung

Im Rahmen der Videoauswertung wurde von Dahmen (2012) für jeden Schüler jeweils ein eigenes Video für beide Experimentierhandlungen (Pipettierübung und Gel beladen) angelegt. Zur Festlegung der Schnittgrenzen wurden Beginn und Ende der Experimentierhandlung explizit definiert. Demnach beginnt eine Experimentierhandlung, wenn der erste und mindestens ein weiterer der in Tabelle 75 (linke Spalte) genannten Punkte erfüllt sind. Die Handlung endet, wenn der erste und mindestens ein weiterer der in Tabelle 75 (rechte Spalte) aufgeführten Punkte eintreten. Der Videoschnitt wurde in Video-Dub (Version 1.8.12) vorgenommen. So entstanden Videosequenzen von etwa einer Minute Länge für jede Experimentierhandlung und jeden Schüler.

Tabelle 75: Definition der Schnittgrenzen (SuS: Schülerin/Schüler; in Anlehnung an Dahmen 2012)

Beginn der Experimentierhandlung	Ende der Experimentierhandlung
<ol style="list-style-type: none"> 1. SuS kommt zu der Station, an der die Experimentierhandlung durchgeführt wird 2. SuS beginnt mit den erforderlichen Geräten zu agieren 3. SuS nimmt die Pipette vom Tisch auf 4. SuS übernimmt die Pipette aus der Hand einer Mitschülerin/eines Mitschülers oder einer betreuenden Person 	<ol style="list-style-type: none"> 1. SuS hat den Arbeitsablauf vollständig ausgeführt 2. SuS verlässt die Station 3. SuS legt die Pipette aus der Hand 4. SuS übergibt die Pipette an eine Mitschülerin/einen Mitschüler oder eine betreuende Person

Nach dem Videoschnitt wurde das Verhalten jeden Schülers im Programm Videograph (Version 4.2.19) kodiert („videographiert“). Unter Kodierung wird hier die Messung der zeitlichen Dauer eines

Verhaltens verstanden. Die Kodierung kann in Videograph synchron zur Videosequenz erfolgen und als *time-sampling* (Verwendung segmentierter Zeitintervalle) oder *event-sampling* (ereignisbezogene Kodierung) vorgenommen werden (IPN Kiel 2012). Aufgrund der erwarteten Kürze der zu beobachtenden Verhaltensweisen wurde hier das *event-sampling* als Methode gewählt. Zur Beurteilung der laborpraktischen Experimentierfähigkeit wurden in der vorliegenden Studie die gemeinsam mit Dahmen (2012) erstellten Messdimensionen Fehler bei der Bedienung von Laborgeräten (Messdimension: F), Fehler in den einzuhaltenden Sicherheitsmaßnahmen (Messdimension: S), Unsicherheit im Arbeitsablauf (U), Experimentbezogenheit des Handelns (H) sowie Experimentbezogenheit der Kommunikation (K) verwendet (Tab. 76) Für die Messdimension Bedienungsfehler sind beispielsweise die Kodierungen 0 (macht gerade einen Fehler) oder 1 (macht gerade keinen Fehler) möglich.

Tabelle 76: Übersicht der Messdimensionen in der Videoanalyse (verändert nach Dahmen 2012)

Messdimension (Kennbuchstabe)	Bedeutung	Kodierung/Beobachtbares Verhalten
Bedienungsfehler (F)	Fehler in der Bedienung der Laborgeräte	0 Fehler 1 Kein Fehler
Sicherheitsfehler (S)	Fehler in den einzuhaltenden Sicherheitsmaßnahmen (z.B. Brille nicht getragen, Handschuhe vergessen...)	0 Fehler 1 Kein Fehler
Unsicherheit (U)	Unsicherheit im Arbeitsablauf	0 unsicher 1 mäßig sicher 2 sicher
Handlung (H)	Handlung experimentbezogen (oder anderweitig orientiert)	0 abgelenkt 1 experimentbezogen
Kommunikation (K)	Kommunikation experimentbezogen (oder anderweitig orientiert)	0 keine Kommunikation 1 Kommunikation nicht experimentorientiert 2 Kommunikation experimentorientiert 3 Kommunikation nicht zu beurteilen

Die objektive Auswertbarkeit der Videos wurde durch den in Tabelle 77 abgebildeten Kodierleitfaden möglich. Dieser formuliert ausführlich, welches Experimentierverhalten mit welcher Kodierung versehen wurde.

Tabelle 77: Kodierleitfaden für die Messdimensionen (verändert nach Dahmen 2012)

Messdimension	Kodierung	Beobachtbares Verhalten
F	0	Der Schüler macht einen Fehler in der Bedienung der Laborgeräte. Im Umgang mit der Pipette kann dies z.B. sein: Pipette auf den Kopf stellen, Pipettenspitze vergessen, Spitze mit den Fingern abnehmen, Druckknopf erst nach Eintauchen in die Flüssigkeit drücken)
F	1	Der Schüler macht keinen der unter 0 aufgeführten Fehler in der Bedienung der Laborgeräte.
S	0	Der Schüler macht einen Fehler in Bezug auf die Sicherheitsvorkehrungen. Hauptsächlich fällt hierunter das Vergessen von Handschuhen oder Schutzbrille. Da die Schüler bei der Pipettierübung nur Wasser pipettieren, wird ihnen von den Betreuern gesagt, Handschuhe und Brille sind hier noch nicht notwendig. Insofern können Sicherheitsfehler bei der Pipettierübung nicht kodiert werden.
S	1	Der Schüler macht keine sicherheitsrelevanten Fehler, er trägt Schutzbrille und Handschuhe und befolgt auch sonst alle zuvor bekannt gegebenen Sicherheitsregeln.
U	0	Der Schüler ist unsicher. Zur Beurteilung der Unsicherheit werden die Kriterien nach Mayring et al. (2005) zu Lernemotionen herangezogen. Da auf den Videos die Mimik oft nur schlecht zu erkennen ist, wird dieser Punkt hauptsächlich anhand von Gestik/Handlungstendenz und verbalen Äußerungen bewertet. Als Zeichen von Unsicherheit werden gewertet: unsichere, verkrampfte Körperhaltung, Anspannung, häufiges Nachfragen, ob alles richtig ist, ständiger Versuch Blickkontakt zum Mitschüler oder Betreuer herzustellen um sich rückzuversichern, zögerliches Vorgehen, häufiges Unterbrechen des Arbeitsablaufes zur Selbst- oder Fremdkontrolle (vom Schüler ausgehend) oder um das Laborgerät noch einmal genauer zu betrachten (z.B. wird nach dem Abwurfknopf für Pipettenspitzen gesucht oder immer wieder das eingestellte Volumen überprüft). In vielen Fällen kann Sicherheit oder Unsicherheit über den kompletten Arbeitsablauf mit einem Wert codiert werden. Da die Videos recht kurz sind, tritt meist kaum eine Veränderung in diesen Minuten ein. Eine Ausnahme kann ein unsicherer Schüler sein, der durch Feedback oder Lob durch Betreuer oder Mitschüler im folgenden Arbeitsablauf eher sicher wird.
U	1	Der Schüler ist mäßig sicher. Das Verhalten des Schülers ist weder als eindeutig sicher, noch als eindeutig unsicher zu kategorisieren.
U	2	Der Schüler ist sicher. Als Zeichen von Sicherheit werden gewertet: lockere Körperhaltung, entspannter Gesichtsausdruck, zügiger Arbeitsablauf mit wenigen Unterbrechungen die vom Schüler ausgehen, wenig bis kein rückversichernder Blickkontakt. Nachfragen sind selten, wenn welche gestellt werden sind sie eher inhaltlich orientiert als rückversichernd zum eigenen Handeln. Die Funktionen des Laborgerätes werden ohne nochmaliges überlegendes Betrachten genutzt (z.B. kann die Spitze ohne nochmaliges Betrachten der Pipette korrekt abgeworfen werden). In vielen Fällen kann Sicherheit oder Unsicherheit über den kompletten Arbeitsablauf mit einem Wert codiert werden. Da die Videos recht kurz sind, tritt meist kaum eine Veränderung in diesen Minuten ein. Eine Ausnahme kann ein sicherer Schüler sein, der durch Anweisungen, Nachfragen oder Fehlerkorrektur durch Betreuer oder Mitschüler unsicher wird.

Messdimension	Kodierung	Beobachtbares Verhalten
H	0	Die Handlung des Schülers ist abgelenkt. Der Schüler zeigt keine der unter 1 genannten Handlungen.
H	1	Die Handlungen des Schülers sind experimentbezogen. Sie/Er ist dem Versuchsaufbau zugewandt und/oder beschäftigen sich mit der Handhabung von Versuchsmaterialien. Auch Weitergabe von Versuchsmaterialien an Nachbargruppen/den Betreuer oder Empfang von Versuchsmaterialien von Nachbargruppen/dem Betreuer wird als handlungsbezogen klassifiziert.
K	0	Der Schüler kommuniziert nicht.
K	1	Der Schüler kommuniziert über Dinge, die nichts mit dem Experiment oder dem gerade laufenden Arbeitsablauf zu tun haben (z.B. über private Dinge).
K	2	Der Schüler kommuniziert über Dinge, die unmittelbar mit dem Arbeitsablauf oder dem Experiment zusammenhängen. Inhaltliche Nachfragen zum Themenfeld des Kurses werden auch als experimentbezogen klassifiziert.
K	3	Die Kommunikation ist nicht zu beurteilen. Mögliche Gründe: schlechte Tonqualität, Störgeräusche etc.

Die Zeit, die die Schüler zur Durchführung der Pipettierübung und zum Beladen des Gels benötigten, variierte personenabhängig stark. Um eine Vergleichbarkeit der Daten zu gewährleisten, wurden die kodierten Verhaltensdauern daher in prozentuale Anteile der Gesamtdauer der Tätigkeit umgerechnet. Diese Prozentwerte wurden im Rahmen der Hauptstudie eingesetzt, um die tatsächliche, im Labor bewiesene Experimentierfähigkeit der Schüler mit der Selbsteinschätzung im FSWEx abzugleichen. Als Korrelationskoeffizient wurde Kennalls Tau berechnet.

3 Ergebnisse

Die Ergebnisdarstellung gliedert sich in einen fachwissenschaftlichen und einen fachdidaktischen Abschnitt. Im fachwissenschaftlichen Teil werden die Resultate der laborpraktischen Optimierungen für die Schülerlaborkurse Genetischer Fingerabdruck und LDH berichtet. Der fachdidaktische Abschnitt geht zunächst auf den im Rahmen dieser Arbeit neu entwickelten Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeitserwartung (FSWEx) ein. Abschließend folgt die Darstellung der zentralen Ergebnisse der Haupt- und Nebenstudie, sowie der Befragung der Lehrerinnen und Lehrer.

3.1 Fachwissenschaftliche Ergebnisse

3.1.1 Optimierung des Genetischen Fingerabdrucks

Zunächst werden in diesem Abschnitt die verschiedenen DNA-Extraktionsmethoden, anschließend die optimierten PCR-Bedingungen zur Erzielung eindeutiger Bandenmuster beim Genetischen Fingerprinting, sowie für Schülerlaborbedingungen geeignete Färbemöglichkeiten der DNA-Banden vorgestellt.

3.1.1.1 DNA-Extraktion mit dem DNeasy Plant Mini Kit, der Chelex-Extraktion und der Kalium-Acetat-Isopropanol-Fällung im Vergleich

Die aus den Buccalzellen von freiwilligen Spendern gewonnenen DNA-Proben wurden mit drei verschiedenen Extraktionsmethoden isoliert. Ziel dieser Untersuchungen war es, eine effektive, kostengünstige und ungefährliche Methode zu finden. Um die Ergebnisse der drei Extraktionsmethoden vergleichen zu können, wurden die extrahierten DNAs in mehreren PCR-Reaktionen mit Primern für den VNTR-Locus D1S80 (Tab. 7) verwendet, die unter identischen Bedingungen durchgeführt wurden (Tab. 9 und 13). Dabei ist zu erkennen (Abb. 37), dass sowohl die Extraktion mit dem DNeasy-Kit (Spuren 1 und 2), als auch die Chelex-Extraktion (Spuren 3 und 4) sehr gute Resultate erbrachten. Die bei allen Probanden zu detektierende Bande bei ca. 500-540 bp ist auch bei allen Extraktionsmethoden angefärbt. Allerdings sind die Banden bei der Kaliumacetat-Isopropanolfällung (Spuren 5-7) schwächer ausgeprägt, aber immer noch sichtbar. Für den Einsatz im Schülerlabor wurde daher die relativ kostengünstige Chelex-Extraktion zur Vorbereitung der DNA-Proben durch die Laborkursbetreuer gewählt. Diese DNA-Proben wurden dann von den Schülern im Schülerlaborkurs als Tatverdächtigen-DNA verwendet. Für die von den Schülern durchzuführende Extraktion wurde trotz geringerer Effektivität die noch kostengünstigere Kaliumacetat-

Isopropanolfällung umgesetzt. Für diese wurde – wie in Tab. 6 dargestellt – eine zeitökonomische Version gefunden.

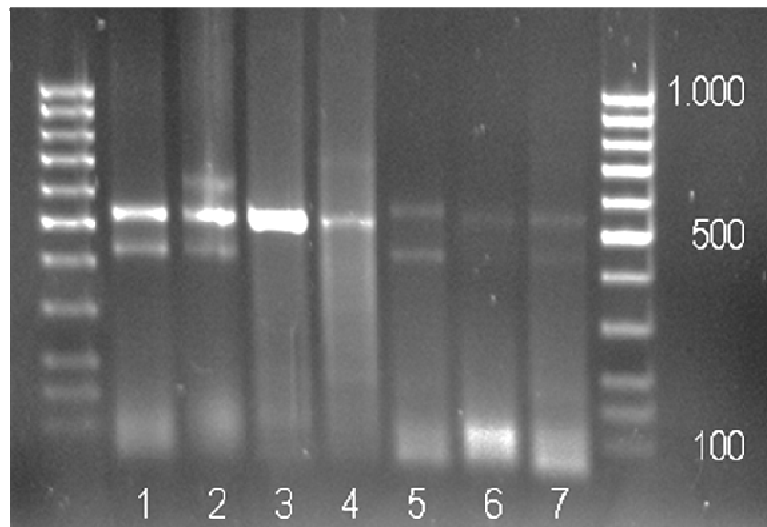


Abbildung 37: Gelelektrophoretische Auftrennung der PCR-Produkte des VNTR-Locus D1S80 nach verschiedenen DNA-Extraktionsmethoden (Agarose 1,5% w/v). Rechts und links 100 bp-DNA-Ladder Spuren 1 + 2: DNeasy-Extraktion; Spuren 3+4: Chelex-Extraktion; Spuren 5-7: Kaliumacetat-Isopropanol-Fällung. Färbung Stain G.

Demnach erfüllen das Chelex-Extraktionsverfahren sowie die Kaliumacetat-Isopropanolfällung alle vorher gestellten Bedingungen der hinreichenden Effektivität, Kostengünstigkeit und Ungefährlichkeit für einen Einsatz im Schülerlabor.

3.1.1.2 Sichtbarmachung der DNA-Banden im Agarose Gel

Auch für die Anfärbung der DNA-Banden im Gel sollte ein zuverlässiger und ungiftiger Farbstoff gefunden werden. Um zu überprüfen, ob auch relativ schwache DNA-Banden mit weniger carcinogenen Methoden als Ethidiumbromid sichtbar gemacht werden können, wurden Gele mit Ladder-DNA (0,1µg DNA/µL) in identischen DNA-Mengen (von links nach rechts in Abb. 38 Spur 1: 0,05µg Ladder-DNA, Spur 2: 0,1µg, Spur 3: 0,2µg, Spur 4: 0,3µg, Spur 5: 0,4µg, Spur 6: 0,5µg) beladen. Die Gele wurden mit den im UV-Licht fluoreszierenden Farbstoffen Ethidiumbromid (prestain), Gel Red (prestain), Stain G (prestain) und dem nicht-fluoreszierenden Farbstoff Methylblau (poststain) angefärbt (Tab. 15, Abb. 38). Die beiden fluoreszierenden Farbstoffe Stain G (B) und Gel Red (C) färbten die Banden ebenso wie beim Ethidiumbromid (A) bereits bei niedrigen Konzentrationen von 0,05µg eingesetzter DNA Ladder relativ gut an. Der nicht-fluoreszierende Farbstoff Methylblau (D) zeigte deutlich schlechtere Ergebnisse. Relativ deutlich zu sehen sind die Banden hier erst bei 0,2µg aufgetragener DNA-Ladder. Insbesondere die Banden kurzer DNA-Fragmente sind hier, wie auch bei den höheren Konzentrationen, kaum zu erkennen. Da die

erwarteten Banden der VNTR-Loci häufig im unteren Drittel der Leiter erwartet werden, sind nur solche Farbstoffe geeignet, die auch diese gut detektieren, so dass Methylenblau für die Verwendung im Schülerlabor ausgeschlossen wurde. Ein weiterer Nachteil der nachträglichen Färbung mit Methylenblau war der – im Vergleich zu den UV-Farbstoffen – zusätzliche Zeitaufwand für die Färbung und die anschließende Entfärbung.

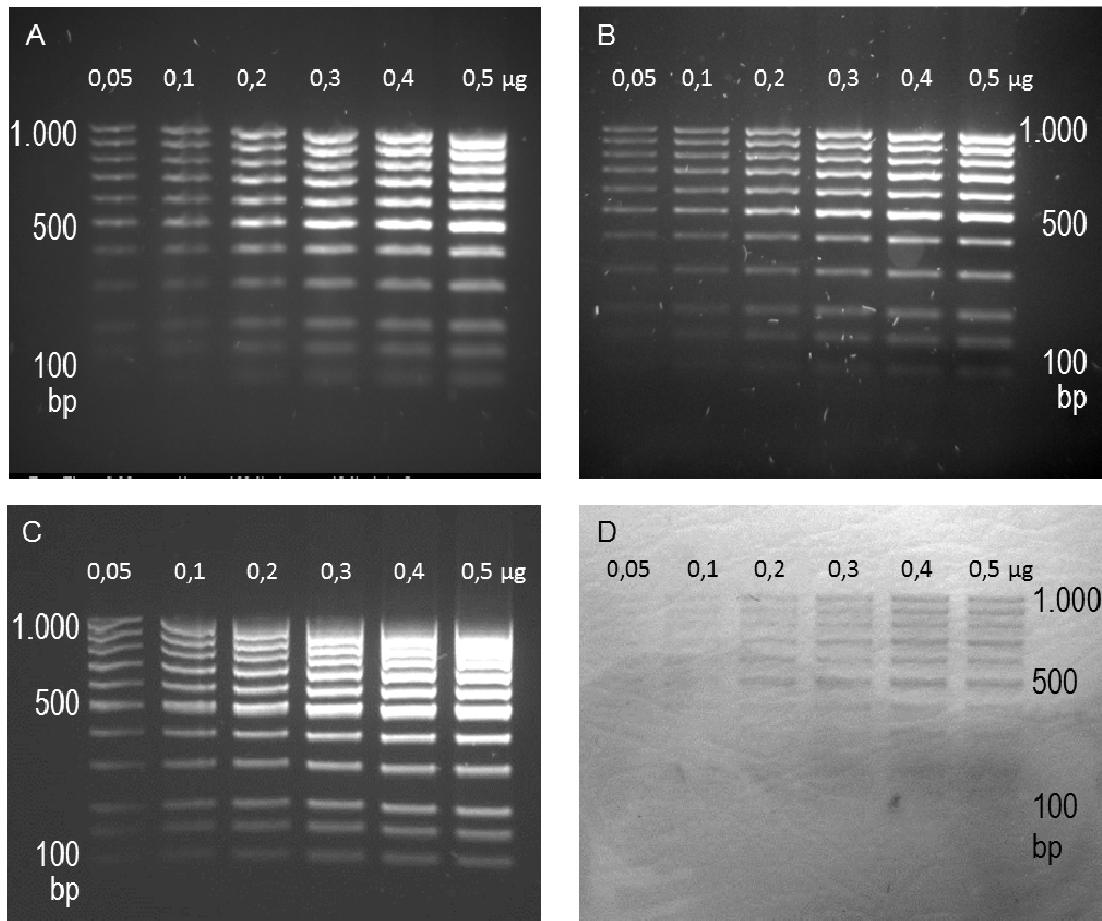


Abbildung 38: Gelelektrophoretische Auftrennung (Agarose 1,5% w/v) des Markers 100 bp-DNA-Ladder equimolar mit verschiedenen Färbemethoden in folgenden Konzentrationen: Spur 1: 0,05 µg; Spur 2: 0,1 µg; Spur 3: 0,2 µg; Spur 4: 0,3 µg; Spur 5: 0,4 µg; Spur 6: 0,5 µg. Anfärbung A: Ethidiumbromid; B: Stain G; Gel Red; Methylenblau. A-C unter UV-Bestrahlung bei 312 nm.

Für die Verwendung im Schülerlabor sowie für die nachfolgend vorgestellten Experimente wurde Stain G als nachweisempfindliche, relativ ungefährliche (im Vergleich zu Ethidiumbromid) und relativ kostengünstige Variante ausgewählt.

3.1.1.3 Untersuchungen zur Optimierung der PCR-Bedingungen

Nachdem die beiden o.g. DNA-Extraktionsmethoden sowie ein Verfahren zur DNA-Färbung für das Schülerlabor favorisiert wurden, mussten nun die Bedingungen der PCR zur Analyse der VNTRs aus den Mundschleimhautzellen untersucht werden. Auch hier galt es, bestimmte Ziele zu erreichen.

Dazu gehörten Effektivität der eingesetzten Primer, hohe Reproduzierbarkeit der Ergebnisse, klar abgrenzbare Allele der untersuchten DNA-Abschnitte, gute Nachweisbarkeit aller in einer Schulklasse vorkommenden Allele, Ungefährlichkeit und Kostengünstigkeit. Die Annealingtemperatur, also die Temperatur, bei der sich die DNA-Primer an die DNA-Vorlage (und später an die bereits hergestellte neue DNA) anlagert, damit die Taq-Polymerase den Gegenstrang komplementieren kann, stellt einen sehr sensiblen Faktor für eine erfolgreiche Amplifikation der DNA dar. Daher wurde in einem ersten Schritt eine Gradienten-PCR durchgeführt, bei der verschiedene Temperaturen in einem PCR-Lauf, also unter identischen Bedingungen, erprobt werden konnten. Hier dargestellt ist exemplarisch das Agarose-Gelbild einer Gradienten-PCR mit der Primerkombination zum VNTR-Locus D17S30 (Abb. 39).

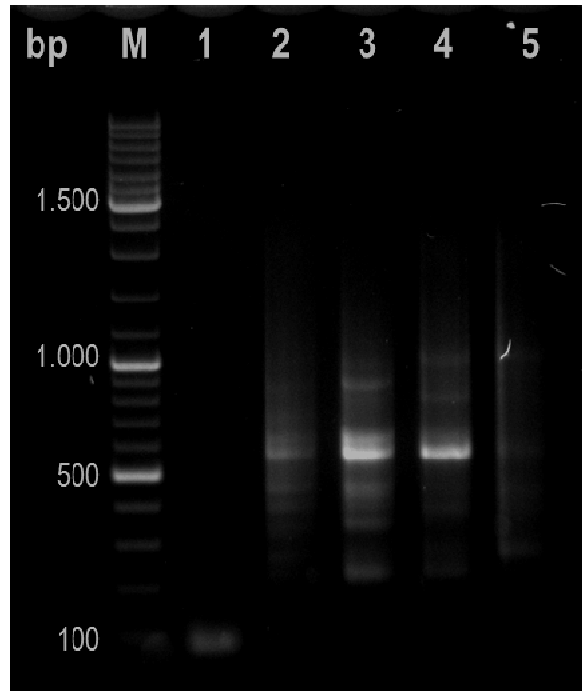


Abbildung 39: Gel-Elektrophoretische Auftrennung der PCR-Produkte nach Gradienten-PCR mit dem Primerpaar für den Locus D17S80. (Agarose 1,5% w/v) bei 312 nm, M: Marker 100 bp-DNA-Ladder. Spur 1: 45°C; Spur 2: 50,4°C; Spur 3: 55,8°C; Spur 4: 61°C; Spur 5: 65,6°C; 5 µl PCR-Produkt je Spur.

Es ist zu sehen, dass die PCR sowohl in den Spuren 2, 3 und 4 eine relativ deutliche Bande bei ca. 560 bp ergibt. Allerdings treten sowohl in Spur 2

(50,4°C), als auch in Spur 3 (55,8°C) und Spur 4 (61,0 °C) unspezifische Nebenbanden auf. Niedrigere (Spur 1; 45°C) oder höhere Temperaturen 65,6°C (Spur 5) zeigen keine deutlich identifizierbaren Banden. In jeder Spur sollten nur eine (homozygot) oder maximal zwei Banden (heterozygot) auftreten. Aufgrund der relativ deutlichen Bande bei 560 bp in Spur 3 (55,8°C) wurde diese als beste Lösung erachtet. Wie die unspezifischen Banden zeigen, konnte allerdings für diesen VNTR-Locus, trotz der sowohl berechneten als auch experimentell erhobenen Annealingtemperatur, kein ideales Ergebnis erzielt werden. In der Diskussion wird darauf später noch einmal eingegangen werden.

Nachdem die idealsten Annealing-Temperaturen für die Primer-Kombinationen ermittelt wurden, mussten noch weitere PCR-Bedingungen angepasst werden. Dazu gehören die Anzahl der Zyklen, die zyklische Denaturierung, sowie die Zeiten für Elongation und die initiale Denaturierung (Tab 10-13). Außerdem wurde die Konzentration an eingesetzter DNA variiert (Tab. 9). Als idealste Lösung stellten

sich die in Tab. 78 für die jeweiligen DNA-Abschnitte dargestellten PCR-Bedingungen heraus. Mit ihnen wurde für jeden Locus ein spezifischer Genetischer Fingerabdruck erstellt, wie er auch im Schülerlabor umgesetzt werden sollte. Das Ziel war es, einen Locus zu identifizieren, der unter den Bedingungen des Schülerlabors eine zuverlässige Zuordnung zulässt.

Tabelle 78: PCR-Bedingungen für das Schülerlabor optimiert

PCR – Protokoll	Temp. [°C] Zeit[min]	Temp. [°C] Zeit[min]	Temp. [°C] Zeit[min]	Temp. [°C] Zeit[min]
VNTR-Locus	D17S30	apoB3	I_g - J_H	D1S80
Initialdenaturierung	95 5	95 5	95 5	95 5
Amplifizierungszyklen	30 x	30 x	35 x	30 x
Denaturierung	94 1	94 1	94 1	94 1
Annealing	56 2	51 1	65 1	65 1
Elongation	72 2	72 2	72 1	72 2
Finale Elongation	72 10	72 10	72 10	72 10

Nachdem nun die Bedingungen für eine vermutlich erfolgreiche Amplifizierung der gesuchten DNA-Abschnitte aufgezeigt wurden, wurde nun überprüft, ob der Nachweis auch unabhängig vom untersuchten Allel funktionierte. In Abhängigkeit von der ethnischen Zugehörigkeit kann die Häufigkeit des Auftretens bestimmter Allele variieren (Deka et al. 1992a/b, Das & Mastana 2003). Um die Wahrscheinlichkeit zu erhöhen, trotz der kleinen untersuchten Stichprobe möglichst viele verschiedene Allele abzudecken, wurden alle VNTR-Loci mit den geeigneten Primerkombinationen und PCR-Bedingungen an freiwilligen Speichelspendern verschiedener ethnischer Herkunft getestet. Dieses Vorgehen diente weder der Zuordnung zu einer bestimmten Ethnie, noch der Identifizierung derselben, sondern diente lediglich der Abdeckung eines breiten Allelspektrums in der untersuchten Stichprobe. Diese Untersuchungen dienten auch nicht der quantitativen Bestimmung von Allelfrequenz bei verschiedenen Ethnien. Sie sollten lediglich der qualitativen (da statistisch nicht ausreichenden) Überprüfung der Annahme dienen, dass die gewählten DNA-Abschnitte unabhängig vom Allel methodisch detektierbar sind. Es wurden deshalb Buccalschleimhautproben von mehreren Probanden verschiedener ethnischer Zugehörigkeit untersucht. Darunter waren als Kaukasier zwei Mitteleuropäer (ME) und ein Osteuropäer (OE). Aus der asiatischen Ethnie wurden je

ein Proband aus Syrien (Syr) und Indien (Ind) getestet. Von der afrikanischen Ethnie stand ein Proband aus Marokko (Mar) zur Verfügung. Die gewonnene DNA wurde für die Amplifizierung der vier untersuchten VNTR-Loci verwendet. Zur besseren Vergleichbarkeit sind die Ergebnisse nach Locus aufgeteilt für jeweils alle Probanden dargestellt (Abb. 40).

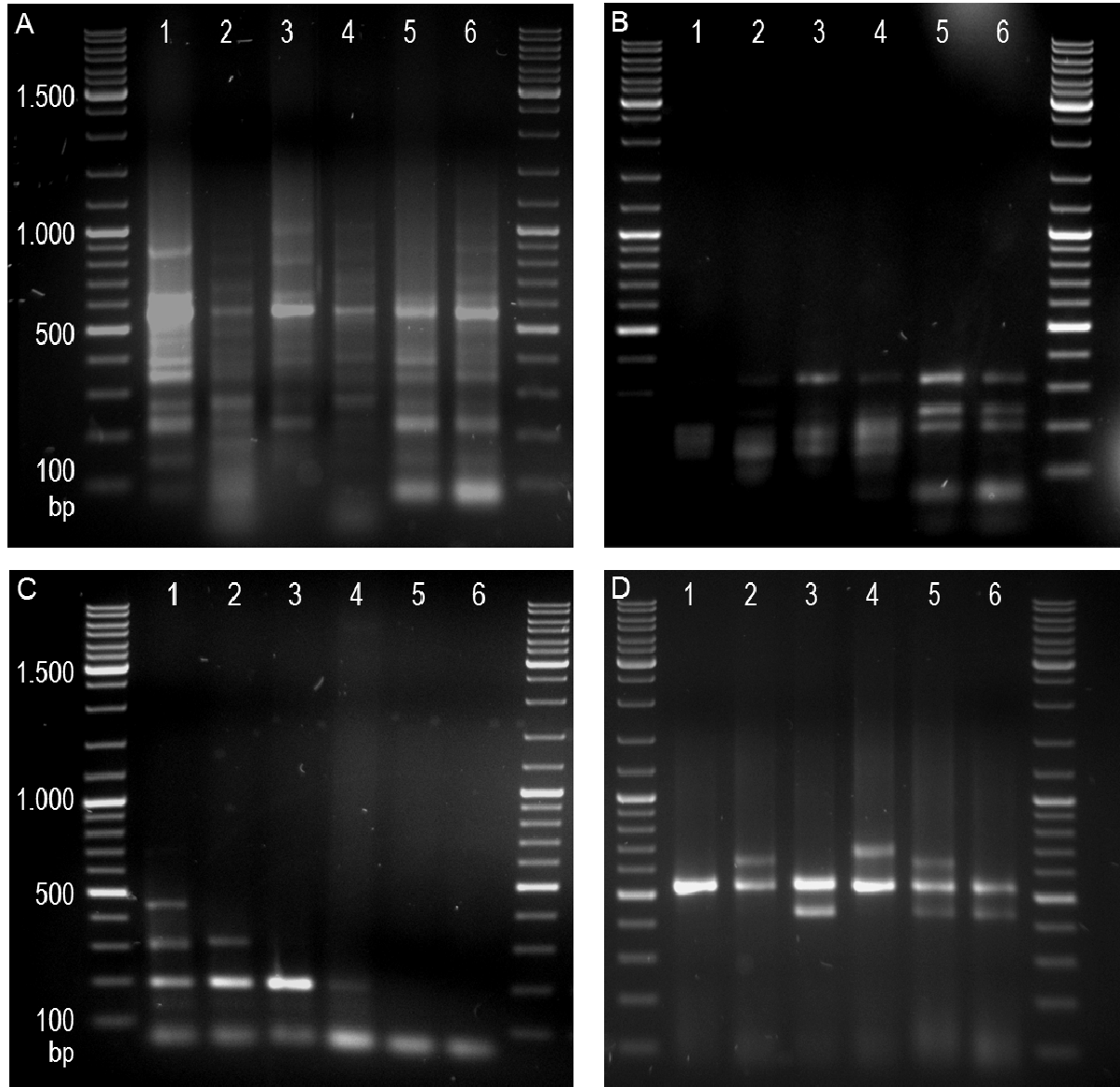


Abbildung 40: Gelelektrophoretische Aufnahme (Agarose 1,5% w/v) der untersuchten VNTR-Loci nach PCR mit Proben von Spendern unterschiedlicher ethnischer Herkunft. A: VNTR-Locus D17S30, B: VNTR-Locus I_G-J_H, C: VNTR-Locus apoB3, D: VNTR-Locus D1S80. Spuren auf jedem Gel identisch: Spur 1: ME; Spur 2: ME; Spur 3: Syr; Spur 4: OE; Spur 5: Mar; Spur 6: Ind. Rechts und links 1 kb Marker DNA-Ladder Mix [bp].

3.1.1.3.1 VNTR-Locus D17S30

Der genetische Fingerabdruck der drei untersuchten ethnischen Gruppierungen dieses Locus ergibt eine Vielzahl von Banden (Abb. 40 A). Bisher sind insgesamt 15 Allele, d.h. 15 verschiedene Sequenzwiederholungen, dieser 70 Basenpaar langen Wiederholungssequenzen bekannt (Deka et al.

1992b). Die Länge der detektierten Banden ergibt sich aus den flankierenden Primern und der Sequenz. In allen Spuren kommen PCR-Produkte mit ca. 560 bp vor. Allerdings treten bei allen Probanden zahlreiche unspezifische Banden auf, so dass eine gezielte Amplifizierung des VNTR Locus D17S30 nicht erfolgreich umgesetzt werden konnte. Zwar unterscheiden sich die Bandenmuster im Gel zwischen den Probanden, dennoch ist eine eindeutige, auf den VNTR D17S30 begründete, Unterscheidung von Personen nicht gegeben. Entsprechend ist auch eine Zuordnung zur Heterozygotie oder zur Homozygotie als wichtiger Parameter beim genetischen Fingerabdruck nicht möglich. Unter den verwendeten PCR-Bedingungen war dieser Locus daher für einen Einsatz im Schülerlabor nicht geeignet.

3.1.1.3.2 VNTR-Locus I_G-J_H

Dieser DNA-Abschnitt zeichnet sich durch Wiederholungssequenzen von 50 bp aus (Abb. 40 B). Sechs Allele wurden von Decorte et al. (1990) identifiziert. Die erwarteten Größen von 470 bis 920 bp (Decorte et al. 1990) konnten in vorliegender Studie nicht reproduziert werden. Auch durch Variationen der PCR-Bedingungen wurden immer die gleichen Bandenmuster erzielt. Da hier bei den meisten getesteten Personen mehr als die zwei maximal erwarteten Banden auftraten und diese zudem eine nicht erwartete Größe aufwiesen, erscheint auch dieser Locus – unter den getesteten PCR-Bedingungen – für den Einsatz im Schülerlabor ungeeignet.

3.1.1.3.3 VNTR-Locus apoB3

Dieser nahe dem Signal für das Transkriptionsende des Apolipoprotein B gelegene Locus besitzt eine Core-Sequenz von 14-16 Basenpaaren (Das & Mastana 2003). Zu erwarten ist eine Länge zwischen 300 und 900bp (vgl. Abschnitt 1.1.1). Damit liegen in Abbildung 40 C die Banden zum Teil unterhalb der erwarteten Größenordnung. Bei den Personen 5 und 6 treten gar keine Banden – mit Ausnahme der Primerbanden – und bei Person 1 bis 4 eine zusätzliche unerwartete Bande bei ca. 200bp auf. Weiterhin liegen die erzielten Banden bei unterschiedlichen Personen auffälligerweise auf gleicher Höhe. Da eine sichere Unterscheidung zwischen zwei Individuen nicht möglich ist, wird auch dieser Locus unter den getesteten Bedingungen als ungeeignet für den Einsatz im Schülerlabor bewertet.

3.1.1.3.4 VNTR-Locus D1S80

In diesem auf dem Chromosom 1 gelegenen Locus kommen (mindestens) vier der bisher beschriebenen 32 verschiedenen Allele vor. Sämtliche detektierten Banden liegen im erwarteten Größenbereich von ca. 200-700bp (vgl. Abschnitt 1.1.1). In Spur 1 finden wir einen homozygoten Genotyp bei ca. 530 bp. Eine Bande in ähnlicher Größe ist auch in den anderen Proben aufzufinden. In Spur 2 findet sich zusätzlich noch eine Bande bei 640 bp. Spur 3 zeigt ebenso wie die Spuren 5 und 6 eine Bande bei 430 bp. In Spur 4 kommt neben der 530 bp-Bande noch eine deutliche bei ca. 700 bp heraus. Bei den Spur 5 ist nicht sicher abzuklären, ob die bei ca. 620 bp liegende oder die bei 430 bp liegende Bande unspezifisch ist. Für diese sehr kleine Stichprobe von N= 6 ist die Verteilung der Allele sehr hoch (Abb. 40 D). Die Wiederholungseinheiten sind hier mit 16 bp recht klein und evtl. zwei ähnlich lange Allele schwierig zu unterscheiden. Lediglich in Spur 3 (Syr) und 6 (Ind) kommen identische Genotypen vor. Damit ist gewährleistet, dass mit diesem Marker, unabhängig von der ethnischen Zugehörigkeit (bei aller hier fehlenden Statistik), mit hoher Wahrscheinlichkeit Unterschiede im Genotyp festgestellt werden können. Die beiden Mitteleuropäer unterscheiden sich in zwei Genotypen, wobei in Spur 1 mit nur einer Bande im Bereich von ca. 530 bp ein homozygotes Merkmal detektiert wurde. In der vorliegenden Stichprobe kann also sowohl eine diskrete Allel-Unterscheidung vorgenommen werden, als auch zwischen homozygoten (Spur 1) und heterozygoten Personen (Spur 2) unterschieden werden. Da dieser VNTR-Locus die Bedingung hoher Heterogenität erfüllt, unter den erprobten PCR-Bedingungen distinkte Allele zeigt und eine relativ gute Unterscheidung zwischen homo- und heterozygoten Probanden zulässt, wird er als geeignet für die Schülerlaborkurse angesehen.

3.1.1.3.5 Die VNTR-Loci im Vergleich

Dieses wird sehr deutlich, wenn man alle PCR-Produkte der VNTRs exemplarisch für einen Kaukasier (ME, jeweils die ungerade Spur) und für einen Asiaten (Syr, jeweils die gerade Spur) auf ein Agarose-Gel aufträgt und vergleicht (Abb. 41). Eine eindeutige Unterscheidung ist nur in den Bandenmustern des Locus D1S80 in den Spuren 3 (ME) und 4 (Syr) möglich.

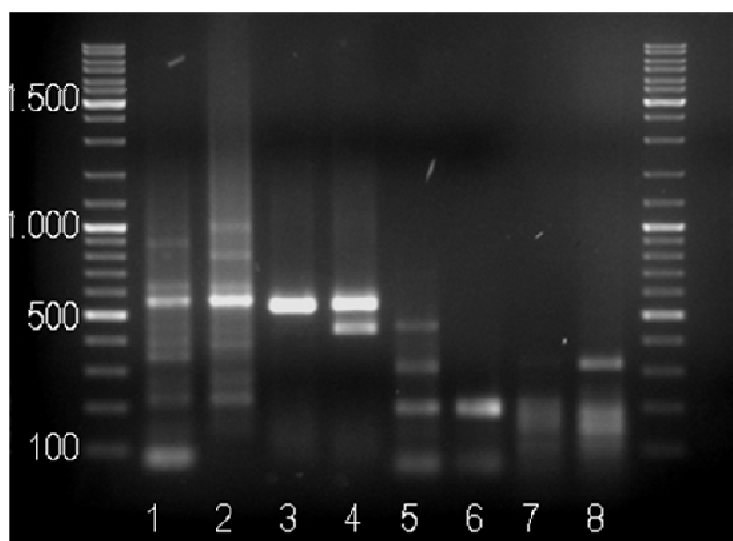


Abbildung 41: Gelelektrophoretische Aufnahme der PCR-Produkte aller vier VNTR-Loci für ME (Spuren 1, 3, 5, 7) und Syr (Spuren 2, 4, 6, 8). VNTR-Loci: Spuren 1+2: D17S30; Spuren 3+4: D1S80; Spuren 5+6: apoB; Spuren 7+8: JH-IG. 1kb DNA ladder Mix.

Mit diesen molekularbiologischen Untersuchungen und Experimenten zur Optimierung für einen Einsatz im BeLL Bio konnte gezeigt werden, dass sich die VNTR-Loci als Minisatelliten-DNA dazu eignen, Individuen zu unterscheiden. Es sei noch einmal betont, dass es hier lediglich um die Unterscheidung von (letztlich) anonymen Individuen geht und nicht um die Zuordnung einer DNA-Probe zu einem speziellen Individuum oder zu einer ethnischen Gruppierung. Als geeignet erwiesen sich zusammenfassend dargestellt die folgenden Bedingungen:

- Die Kaliumacetat-Isopropanol-Fällung konnte als eine geeignete, kostengünstige und zeitsparende Methode identifiziert werden, um im Schülerlabor von den Schülern selbst DNA aus Mundschleimhautzellen isolieren zu lassen.
- Für die durch das BeLL Bio vorbereiteten DNA-Proben sollte in jedem Fall eine genügend hohe Konzentration an DNA in den Eluaten vorliegen. Dafür wurde die Chelex-Methode ausgewählt. Sie ist effektiver, aber teurer als die Kaliumacetat-Isopropanol-Fällung.
- Die Färbung mit Stain G als Fluoreszenzfarbstoff, der bei 312 nm detektiert werden kann, erwies sich als die geeignete, relativ ungefährliche und zuverlässige Methode im Schülerlabor.
- Für die Optimierungen der PCR-Bedingungen wurde eine Gradienten-PCR durchgeführt, um die Annealing-Temperatur experimentell zu testen. Es konnte zwar für jedes Primerpaar eine optimale Temperatur ermittelt werden, es waren aber dennoch nicht alle Produkte spezifisch. Die besten getesteten Variation flossen in das aktuelle PCR-Protokoll ein.
- Der Locus D1S80 erwies sich unter den getesteten VNTR-Loci als geeignet, um auch unter Schülerlaborbedingungen gute Ergebnisse erwarten zu lassen. Er zeigt eine gute Unterscheidbarkeit von homozygoten und heterozygoten Trägern sowie bei den beprobten Personen (N=6; nicht repräsentativ) eine hohe Allelheterogenität.

3.1.2 Photometrische Messungen zum Pyruvatabbau durch die Laktatdehydrogenase

Zunächst werden in Abbildung 42 die Spektren der in Form der Erfrischungsgetränke eingesetzten Lebensmittelfarbstoffe E122 (Powerade Wild Cherry, 1:4 Verdünnung, grüne Kurve) und E163 (Powerade Zero, unverdünnt, blaue Kurve), sowie das Spektrum des aus verschiedenen Fruchtsaftkonzentraten zusammengesetzten Punika Classics Rote Früchte (unverdünnt, rote Kurve) aufgezeigt (vgl. Tab. 16). Deutlich wird hier, dass das in seiner Zusammensetzung nicht exakt definierte Fruchtsaftgetränk (rote Kurve) im VIS-Bereich kein klar definiertes Absorptionsmaximum aufweist und die Absorption der Lösung im Wellenlängenbereich $\lambda < 500\text{nm}$ deutlich zunimmt. Im Gegensatz dazu zeigen die Getränke, welche die Lebensmittelfarbstoffe E122 und E163 enthalten, relativ deutlich definierte Maxima bei 528nm (Powerade Zero) und 516nm (Powerade Wild Cherry).

Das Absorptionsspektrum von Powerade Wild Cherry weist außerdem eine Schulter bei etwa 550 nm auf.

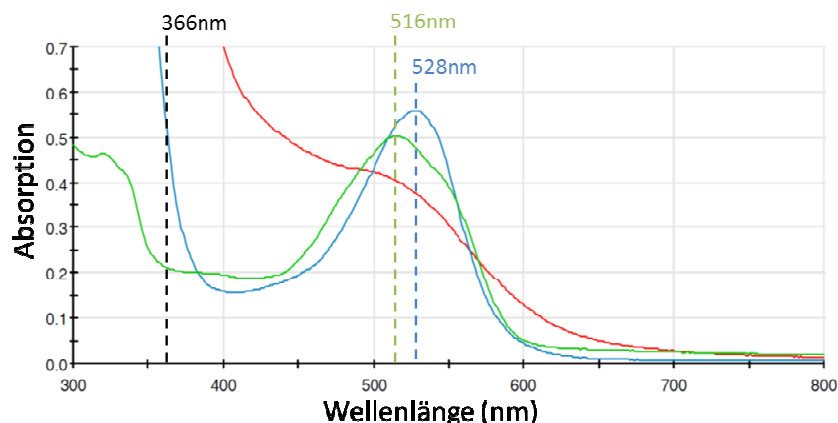


Abbildung 42: Absorptionsspektrum der Erfrischungsgetränke Powerade Sports (grüne Kurve, Absorptionsmaximum: $\lambda=516\text{nm}$; 1:4 Verdünnung des Original-Getränks), Powerade Zero (blaue Kurve, Absorptionsmaximum: $\lambda=528\text{ nm}$; unverdünnt) und Punica Classics Rote Früchte (rote Kurve; unverdünnt) im Wellenlängenbereich von $\lambda=300\text{-}800\text{ nm}$; die NADH-Messung erfolgt bei $\lambda=366\text{ nm}$.

Der NADH-Abbau wird gewöhnlich bei $\lambda=340\text{ nm}$ ($\epsilon_{340}=6,2 \times 10^3\text{ L/mol}\cdot\text{cm}$) oder $\lambda=366\text{ nm}$ ($\epsilon_{366}=3,4 \times 10^3\text{ L/mol}\cdot\text{cm}$) verfolgt. Wie Abbildung 42 zu entnehmen ist, weist in diesem Wellenlängenbereich insbesondere Punica Classics eine sehr hohe Absorption auf. Daher wurden die weiteren Messungen des NADH-Abbaus nur unter Zusatz von Powerade Wild Cherry beziehungsweise Powerade Zero durchgeführt. Die folgende Abbildung 43 zeigt den NADH-Abbau durch die LDH beim Umsatz von Pyruvat zu Laktat für **Probe A** (ohne Zusatz von Powerade).

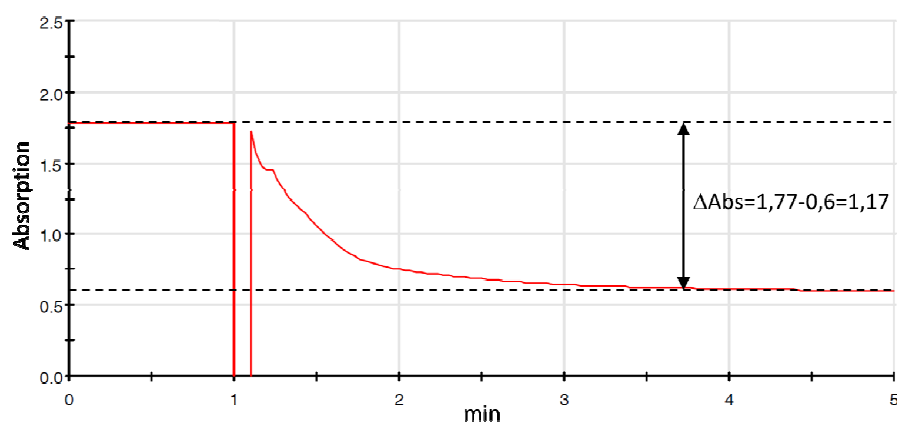


Abbildung 43: Spektrum der Absorptionsänderung in Probe A; Absorptionsmessung von NADH bei $\lambda=366\text{ nm}$ ohne Zusatz von Powerade; Abbau von NADH durch Laktatdehydrogenase (LDH) beim Umsatz von Pyruvat zu Laktat; Zugabe der LDH bei $t=1\text{ min}$

Die photometrische Messung des NADH-Verbrauchs wurde für Probe A fünfmal durchgeführt. Die durchschnittliche Absorptionsänderung beträgt $\Delta\text{Abs}_{\text{A}}=1,17$. Die anhand der Einwaage berechnete Pyruvatkonzentration für die Probe A beträgt $c_{\text{A_Einwaage}}=0,014\text{ mol/L}$. Zur Berechnung der Pyruvatkonzentration aus dem photometrisch verfolgten NADH-Abbau wurde die Differenz von

Ausgangs- und Endabsorption ΔAbs_{\emptyset_A} in das Lambert-Beersche Gesetz eingesetzt. Für Probe A bedeutet dies:

$$c_{A_photometrisch} \left[\frac{mol}{L} \right] = \frac{\Delta Abs_{\emptyset_A}}{\epsilon \left[\frac{L}{mol \times cm} \right] \times d[cm]} = \frac{1,17}{3,4 \times 10^3 \frac{L}{mol \times cm} \times 1 cm} = 3,44 \times 10^{-4} \frac{mol}{L}$$

Unter Berücksichtigung der 40x Verdünnung der Probe A in der Küvette ergibt sich:

$$c_A \left[\frac{mol}{L} \right] = c_{A_photometrisch} \left[\frac{mol}{L} \right] \times 40 = 0,014 \left[\frac{mol}{L} \right]$$

Somit entspricht die aus dem NADH-Abbau berechnete Pyruvat-Konzentration c_A exakt der Konzentration $c_{A_Einwaage}$. Sämtliches Pyruvat in der Lösung wurde somit in etwa fünf Minuten zu Laktat umgesetzt.

Abbildung 44 stellt die Abbaukurve von NADH für **Probe B** (Zusatz von Powerate Wild Cherry) dar.

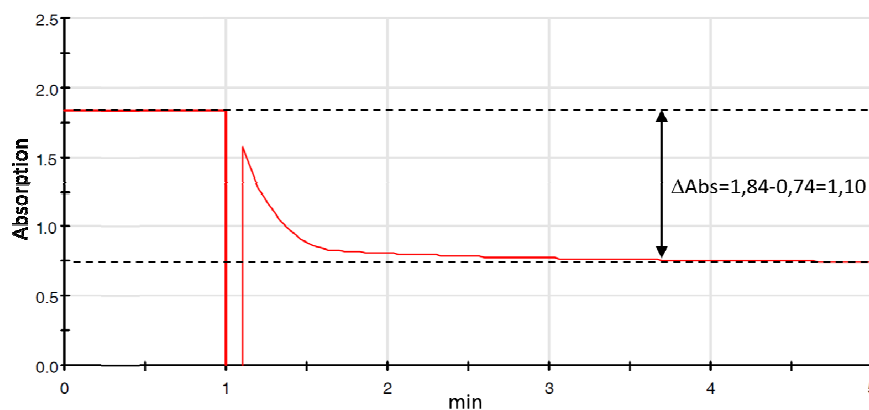


Abbildung 44: Spektrum der Absorptionsänderung in Probe B; Absorptionsmessung von NADH bei $\lambda=366$ nm; 50 μ L Powerade Wild Cherry; Abbau von NADH durch Laktatdehydrogenase (LDH) beim Umsatz von Pyruvat zu Laktat; Zugabe der LDH bei t=1 min

Auch für Probe B wurde die photometrische Messung des NADH-Verbrauchs fünfmal vorgenommen. Die durchschnittliche Absorptionsänderung liegt bei $\Delta Abs_{\emptyset_B}=1,13$. Die anhand der Einwaage berechnete Pyruvatkonzentration für die Probe B beträgt $c_{B_Einwaage}=0,014$ mol/L. Unter Einsatz von ΔAbs_{\emptyset_B} in das Lambert-Beersche Gesetz ergibt sich für Probe B:

$$c_{B_photometrisch} \left[\frac{mol}{L} \right] = \frac{\Delta Abs_{\emptyset_B}}{\epsilon \left[\frac{L}{mol \times cm} \right] \times d[cm]} = \frac{1,13}{3,4 \times 10^3 \frac{L}{mol \times cm} \times 1 cm} = 3,32 \times 10^{-4} \frac{mol}{L}$$

$$c_B \left[\frac{mol}{L} \right] = c_{B_photometrisch} \left[\frac{mol}{L} \right] \times 40 = 0,013 \left[\frac{mol}{L} \right]$$

Für Probe B entspricht die Konzentration c_B annähernd der Konzentration $c_{B_Einwaage}$.

Der Abbau von NADH in **Probe C** ergibt sich aus Abbildung 45 (Zusatz von Powerate Zero).

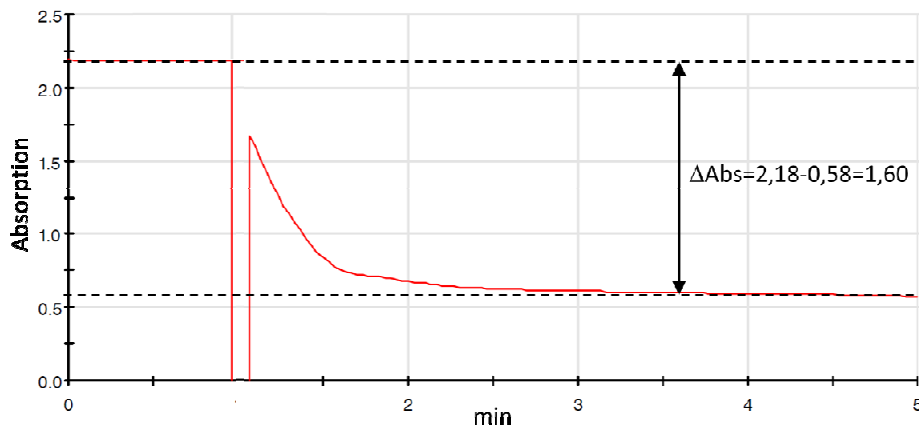


Abbildung 45: Spektrum der Absorptionsänderung in Probe C; Messung von NADH bei $\lambda=366$ nm; Zusatz von 50 μ L Powerade Zero, Abbau von NADH durch Laktatdehydrogenase (LDH) beim Umsatz von Pyruvat zu Laktat; Zugabe der LDH bei t=1 min

Die oben beschriebene Messung wurde für Probe C fünfmal vorgenommen. Anhand der Einwaage wurde für Probe C die Pyruvatkonzentration $c_{C_Einwaage}=0,018$ mol/L berechnet. Unter Einsatz der durchschnittlichen Absorptionsänderung ΔAbs_{\emptyset_C} in das Lambert-Beersche Gesetz ergibt sich für Probe C:

$$c_{C_photometrisch} \left[\frac{mol}{L} \right] = \frac{\Delta Abs_{\emptyset_C}}{\epsilon \left[\frac{L}{mol \times cm} \right] \times d [cm]} = \frac{1,50}{3,4 \times 10^3 \frac{L}{mol \times cm} \times 1 \text{ cm}} = 4,7 \times 10^{-4} \frac{mol}{L}$$

Unter Berücksichtigung der 40x Verdünnung der Probe B in der Küvette ergibt sich:

$$c_C \left[\frac{mol}{L} \right] = c_{C_photometrisch} \left[\frac{mol}{L} \right] \times 40 = 0,018 \left[\frac{mol}{L} \right]$$

Wie bereits bei Probe A und B, so entspricht auch für Probe C die photometrisch bestimmte Pyruvatkonzentration der erwarteten Konzentration $c_{C_Einwaage}$. Für alle untersuchten Proben kann keine bedeutsame Abweichung der photometrisch bestimmten von der tatsächlich eingewogenen Substratkonzentration beobachtet werden. Der Umsatz von Pyruvat zu Laktat ist in allen Proben nach fünf Minuten weitgehend erfolgt. Sowohl Powerade Zero als auch Powerade Wild Cherry erwiesen sich als geeignet, um die Pyruvat-Proben im Schülerlabor rot anzufärben.

Die Optimierungen der biowissenschaftlichen Experimente konnten in die BeLL Bio-Kurse zum Genetischen Fingerabdruck und die photometrische Pyruvatbestimmung einfließen. Die Ergebnisse der fachdidaktischen Evaluierung der entsprechenden Kurse finden sich im folgenden Abschnitt.

3.2 Fachdidaktische Ergebnisse

Da sich ein Teil der später vorgestellten Ergebnisse auf die mittels des neu entwickelten FSWEx-Messinstrumentes erhobenen Daten bezieht, soll dessen Entwicklung hier zunächst vorgestellt werden. Im Rahmen dieser Darstellung werden die Ergebnisse der Testverfahren erläutert, die dazu beitragen sollen, die Qualität des entwickelten Fragebogens zu erfassen. Anschließend erfolgt eine umfassende Darstellung der Ergebnisse von Haupt- und Nebenstudie, sowie der Lehrerbefragung.

3.2.1 Fragebogen zur Erfassung der experimentbezogenen Selbstwirksamkeit (FSWEx)

Der Fragebogen zur Erhebung der experimentbezogenen Selbstwirksamkeit (FSWEx) hebt sich deutlich von den bisher existierenden Messinstrumenten zur Erfassung der Experimentierkompetenz ab. Mit ihm wird nicht versucht, die praktische Experimentierfähigkeit in einem Wissenstest oder in Form aufwändiger praktischer Testverfahren zu erheben. Vielmehr erfasst der FSWEx die subjektive Einschätzung der eigenen Experimentierkompetenz, welche im Weiteren auch als experimentbezogene Selbstwirksamkeit bezeichnet wird. In Anlehnung an vergleichbare Messinstrumente (Wilde et al. 2009) erfolgt die Datenerhebung über eine fünfstufige Likert-Skala mit den Stufen stimmt gar nicht (0), stimmt wenig (1), stimmt teils teils (2), stimmt ziemlich (3) und stimmt völlig (4). Damit orientiert sich das Erhebungsformat des FSWEx an Messinstrumenten wie der Skala zur Erfassung der wahrgenommenen Kompetenz (*perceived competence*) im *Intrinsic Motivation Inventory* von Deci und Ryan (2012).

In den folgenden Abschnitten wird ein Überblick über den Entwicklungsprozess des FSWEx gegeben. Der Entwicklungsprozess für jede der drei Fragebogenversionen beginnt mit der Itemkonstruktion auf Grundlage des Modells experimenteller Kompetenz nach Schreiber et al. (2009). Im Anschluss wird die Fragebogengüte belegt und geprüft, ob die entwickelten Items tatsächlich die drei verschiedenen Messdimensionen Planung, Durchführung und Auswertung von Experimenten erfassen und nicht nur eine gemeinsame Dimension. Abschließend erfolgt eine Prüfung auf Interkorrelationen der drei Subskalen, welche ebenfalls Hinweise liefern, ob die Subskalen des Fragebogens eigenständige Messdimensionen darstellen.

3.2.1.1 Entwicklung der Fragebogenversion FSWEx.V1

3.2.1.1.1 Theoriegeleitete Itemkonstruktion für den FSWEx.V1

Inhaltlich basiert der FSWEx auf dem Modell experimenteller Kompetenz nach Schreiber et al. (2009). In Anlehnung an dieses Modell wurden in einer ersten Kurzversion des FSWEx (FSWEx.V1) für jede

der drei Teilkompetenzen Planung, Durchführung und Auswertung zunächst drei Testitems formuliert. Weil die Testitems dieser Kurzversion noch nicht alle Teilkompetenzen des Modells abdecken, wurde sie in einer späteren Version (FSWEx.V2) um weitere Items ergänzt. Während die beiden genannten Versionen die experimentbezogene Selbstwirksamkeitserwartung situationsunabhängig erfassen, wurde eine dritte Version des FSWEx (postFSWEx.V1) entwickelt, um unmittelbar nach einer Experimentiertätigkeit die gerade erfahrene, situationsgebundene Selbstwirksamkeit zu erfassen.

Tabelle 79: Operationalisierung der Teilkompetenzen Planung, Durchführung und Auswertung im Modell experimenteller Kompetenz (Schreiber et al. 2009) und im FSWEx.V1 (*invers kodiertes Item)

	Teilkompetenz/ Subskalen	Operationalisierung der Teilkompetenzen im Modell experimenteller Kompetenz nach Schreiber et al. 2009	Operationalisierung der Modell- Teilkompetenzen in den Subskalen des FSWEx.V1	Variablen- name
FSWEx.V1	Planung	Fragestellungen entwickeln	Durch Beobachtungen in meinem täglichen Leben fallen mir häufig Fragen ein, die man durch ein Experiment aufklären könnte.	pra11
		Vorgegebene Fragestellungen klären & Versuchsplan entwerfen	Es fällt mir leicht, ein Experiment zur Aufklärung einer vorgegebenen Fragestellung zu entwickeln.	pra27
		Erwartungen Formulieren & Hypothesen bilden	Es gelingt mir, anhand theoretischer Überlegungen, Erwartungen (Hypothesen) über den Ausgang eines Experimentes zu formulieren.	pra16
	Durchführung**	Messungen durchführen	Im Umgang mit Laborgeräten bin ich gut.	pra8
		Messungen durchführen	Für die Durchführung von Experimenten habe ich kein Händchen.*	pra31*
		Messungen dokumentieren	Mit dem Aufschreiben von Versuchsbeobachtungen habe ich Schwierigkeiten.*	pra22*
	Auswertung	Messdaten aufbereiten & Messdaten verarbeiten	Die Auswertung von Messdaten fällt mir leicht.	pra13
		Ergebnis interpretieren	Es ist kein Problem für mich, Versuchsbeobachtungen zu deuten.	pra18
		Ergebnis interpretieren	Ich habe häufig Schwierigkeiten bei der Interpretation von Versuchsergebnissen.*	pra26*

**Die Subskala Durchführung des in Pre- und Follow-up-Test eingesetzten FSWEx.V1 beschränkt sich in der Hauptstudie auf die Operationalisierung der Modellaspekte Messungen durchführen und Messungen dokumentieren. Auf die Operationalisierung der Modellaspekte „Geräte zusammenstellen“ und „Versuchsanordnung aufbauen“ wird verzichtet. Grund hierfür ist, dass der FSWEx.V1 mit dem im Posttest eingesetzten postFSWEx.V1 korreliert werden soll und hier nur die o.g. Modellaspekte berücksichtigt werden können (vgl. 3.2.1.3)

Die theoriebasierte Itemkonstruktion gewährleistete eine möglichst valide Erfassung der Experimentierkompetenz. Tabelle 79 stellt die auf Basis des Modells experimenteller Kompetenz (Schreiber et al. 2009) entwickelten Items des FSWEx.V1 vor.

3.2.1.1.2 Prüfung der Fragebogengüte für den FSWEx.V1

Die *Objektivität* des FSWEx wird durch das – unabhängig von der Person des Auswerter – eindeutig auswertbare Format der Likert-Skala gewährleistet. Zur Bestimmung der *Reliabilität* des FSWEx.V1 wurde dieser zunächst bei den Schülern der Hauptstudie im Pretest eingesetzt. Als Maß für die Reliabilität des Messinstrumentes wurde für jede der drei Subskalen des FSWEx.V1 die interne Konsistenz in Form der Kenngröße Cronbachs α bestimmt (Tabelle 80).

Tabelle 80: Reliabilitäten der Teilkompetenzen des FSWEx.V1 im Pretest

SWEx.V1	Teilkompetenz	Anzahl Items	Reliabilität (Cronbachs α)	N
	Planung	3	0,513	607
	Durchführung*	3 (2)	0,518 (0,576)	608 (609)
	Auswertung	3	0,647	616

*In der finalen Datenauswertung wurde Item pra22 aus der Subskala Durchführung zur Reliabilitätsoptimierung entfernt. Die Reliabilität der Subskala steigt auf 0,576.

Das Item pra22 (Mit dem Aufschreiben von Versuchsbeobachtungen habe ich Schwierigkeiten) wurde in der finalen Auswertung der Hauptstudie aus der Subskala Durchführung entfernt. Die Reliabilität der Subskala konnte dadurch von Cronbachs $\alpha=0,518$ auf Cronbachs $\alpha=0,576$ gesteigert werden. Nach Schön (2007) und Ozkan et al. (2008) können Tests mit einer Reliabilität von Cronbachs $\alpha=0,5$ als reliabel angesehen werden. Eine hinreichende Reliabilität der drei Subskalen des FSWEx.V1 ist somit anzunehmen. Dass der FSWEx.V1 tatsächlich – wie beabsichtigt – die subjektiv wahrgenommene Experimentierkompetenz erhebt, sichert die enge Anbindung an das theoretische Modell experimenteller Kompetenz nach Schreiber et al. (2009). Ein valides Modell vorausgesetzt, so sollte die Validität des FSWEx.V1 hoch sein, wenn es ihm gelingt, das Modell gut abzubilden. Die formulierten Items des FSWEx.V1 bemühen sich, die drei Modell-Teilkompetenzen Planung, Durchführung und Auswertung in den entsprechenden Subskalen des Messinstrumentes zu erfassen. Wie gut ihnen dies gelingt, soll die folgende Faktorenanalyse zeigen.

3.2.1.1.3 Prüfung der Dimensionalität des FSWEx.V1

Für alle Items des FSWEx.V1 wurde eine Matrix der Korrelationskoeffizienten erstellt (N=592). Mit dem Barlett-Test wurde überprüft, wie hoch die Wahrscheinlichkeit ist, dass – trotz beobachteter Korrelationen in der Stichprobe – in der Grundgesamtheit keine Korrelationen zu finden sind (Bühl 2010). Sollte sich dies als wahrscheinlich herausstellen würde dies bedeuten, dass die Items des FSWEx.V1 nur zufällig miteinander korrelieren, aber eigentlich nicht gemeinsam die empfundenen Experimentierkompetenz messen. Eine weitere faktorenanalytische Prüfung wäre somit nicht sinnvoll. Hier bestätigte der Barlett-Test jedoch, dass die Korrelationskoeffizienten sich in der Grundgesamtheit mit hoher Wahrscheinlichkeit von null unterscheiden und somit eine Faktorenanalyse sinnvoll erscheint ($p \leq 0,001$). Auch das Kaiser-Meyer-Olkin-Maß bestätigt mit 0,813 eine hohe gemeinsame Varianz der Variablen. Somit sind die zuvor genannten Variablen sehr gut für eine Faktorenanalyse geeignet (Bühl 2010). Um zu prüfen, ob die Items einer Subskala tatsächlich dasselbe Konstrukt erfassen, wurde eine explorative Faktorenanalyse als Hauptkomponentenanalyse mit Varimax-Rotation durchgeführt. Diese ermöglicht es, eine Sammlung verschiedener Items auf Grund ihrer korrelativen Zusammenhänge zu einer Anzahl überschaubarer und interpretierbarer Faktoren zusammenzufassen (Bühl 2010). Da die Formulierung der FSWEx.V1-Items sich an den drei Subskalen des Modells experimenteller Kompetenz (Schreiber et al. 2009) orientiert, wurden der Faktorenanalyse drei zu extrahierende Faktoren vorgegeben. Idealerweise sollten die Items, welche eine Teilkompetenz abbilden sollen, die höchsten Faktorladungen auf einem gemeinsamen Faktor aufweisen. Alle Items der Subskala Planung sollten also die höchsten Ladungen auf diesem Faktor, dem Faktor Planung aufweisen. Die Ladungen der einzelnen Items auf die entsprechenden Faktoren sind Tabelle 81 zu entnehmen. Einheitlich farblich markiert sind die Items, die ihre höchste Ladung auf einem gemeinsamen Faktor aufweisen. Rot markierte Items gruppieren nicht auf den erwarteten Faktor. Faktorenanalytisch finden sie sich also in einer anderen Subskala wieder, als dies eigentlich bei der Entwicklung der Items beabsichtigt wurde.

Deutlich erkennbar ist, dass die Items pra13, pra18 und pra26 die höchsten Faktorladungen auf dem gemeinsamen Faktor eins aufweisen. Bei diesen drei Items handelt es sich um die Items der Subskala Auswertung. Wie beabsichtigt, scheinen sie sehr gut eine gemeinsame Subskala abzubilden. Weiterhin weisen die Items pra8 und pra31 ihre höchsten Faktorladungen auf dem gemeinsamen Faktor zwei auf. Diese Items wurden formuliert in der Absicht, die Subskala Durchführung des Modells nach Schreiber et al. (2009) abzubilden. Auch dies ist offensichtlich gelungen. Lediglich Item pra22 weist seine höchste Ladung mit 0,575 auf Faktor eins auf und gruppiert damit gemeinsam mit den Items der Subskala Auswertung.

Tabelle 81: Die in der explorativen Faktorenanalyse ermittelten Faktorladungen der FSWEx.V1-Items (Extraktionsmethode: Hauptkomponentenanalyse; Rotationsmethode: Varimax). Einheitlich farblich markiert sind die Items, die ihre höchste Ladung auf einem gemeinsamen Faktor aufweisen. Die rot markierten Items gruppieren anders als erwartet. Bei diesen Items werden auch schwache Faktorladungen aufgeführt (N=592).

		Komponenten		
Item		Faktor 1	Faktor 2	Faktor 3
Items zur Abbildung der Modelldimension Planung	pra11	0,112	0,018	0,702
	pra16	0,755	-0,064	0,329
	pra27	0,236	0,214	0,612
Items zur Abbildung der Modelldimension Durchführung	pra8	0,103	0,690	0,450
	pra22	0,575	0,356	-0,378
	pra31	0,202	0,853	-0,005
Items zur Abbildung der Modelldimension Auswertung	pra13	0,602	0,115	0,311
	pra18	0,706	0,219	0,227
	pra26	0,717	0,182	-0,002

Wie bereits im Rahmen der Reliabilitätsanalyse festgestellt wurde, führt Item pra22 zu einer Verschlechterung der Reliabilität der FSWEx.V1-Subskala Durchführung und wurde daher aus der finalen Auswertung ausgeschlossen. Auch die Items pra11 und pra27, welche entwickelt wurden, um die Subskala Planung zu repräsentieren, laden deutlich auf einen gemeinsamen Faktor. Item pra16 (Es gelingt mir, anhand theoretischer Überlegungen, Erwartungen (Hypothesen) über den Ausgang eines Experimentes zu formulieren) lädt zwar ebenfalls mäßig auf diesen Faktor, viel deutlicher aber mit 0,755 gemeinsam mit den Items der Subskala Auswertung auf Faktor eins. Weil dieses Item inhaltlich besser die Subskala Planung als die Subskala Auswertung im Modell experimenteller Kompetenz repräsentiert und zu keiner Verschlechterung der Reliabilität führt, wird es in der Auswertung in dieser Subskala belassen. Die Faktorenanalyse bestätigt somit, dass die in den Subskalen verbleibenden FSWEx.V1-Items relativ gut drei unterschiedliche Subskalen abbilden.

3.2.1.1.4 Interkorrelationen der Subskalen des FSWEx.V1

Interkorrelationen weisen auf Zusammenhänge zwischen den Subskalen eines Messinstrumentes hin (Mayer et al. 2008). Die Überprüfung der Summenscores der drei Subskalen auf Interkorrelationen soll zeigen, ob diese das Vorliegen von drei eigenständigen Teildimensionen, wie es das Modell von Schreiber et al. (2009) postuliert, unterstützen. Mayer et al. (2008) gehen davon aus, dass mittlere Interkorrelationen auf eigenständige Teildimensionen hindeuten. Zu hohe Interkorrelationen sprechen dagegen eher für ein weniger dimensionales Konstrukt. Die Höhe der bivariaten Korrelationen für die drei Subskalen Planung, Durchführung und Auswertung (Tabelle 82) liegt im

mittleren Bereich, was die Annahme unterstützt, dass die Subskalen eigenständige Unterdimensionen des Konstruktes der experimentbezogenen Selbstwirksamkeitserwartung abbilden.

Tabelle 82: Interkorrelationen der drei Subskalen des FSWEx.V1 im Pretest (N=604)

	Planung	Durchführung	Auswertung
Planung	1	0,308	0,527
Durchführung	0,308	1	0,396
Auswertung	0,527	0,396	1

Sowohl die explorative Faktorenanalyse, als auch die beobachteten Interkorrelationen weisen somit darauf hin, dass die Items des FSWEx.V1 – wie dies auch im Modell experimenteller Kompetenz der Fall ist – drei unterschiedliche Dimensionen abbilden. Unter Berücksichtigung, dass die Items des FSWEx.V1 das Modell Schreibers et al. (2009) noch unvollständig erfassen und in der Literatur häufig erst eine Reliabilität von Cronbachs $\alpha \geq 0,7$ als zufriedenstellend erachtet wird, wurde eine erweiterte Version des FSWEx, der FSWEx.V2, für den Einsatz in der Nebenstudie entwickelt.

3.2.1.2 Entwicklung der Fragebogenversion FSWEx.V2

3.2.1.2.1 Theoriegeleitete Itemkonstruktion für den FSWEx.V2

Der FSWEx.V2 beinhaltet neben den Items des FSWEx.V1 die in Tabelle 83 aufgeführten ergänzenden Items. Der Modellaspekt „Versuchsplan entwerfen“, welcher im Modell Schreibers et al. (2009) nicht eindeutig der Planung oder Durchführung zugeordnet wird, wird im FSWEx.V2 der Subskala Planung zugeordnet. Da der „Umgang mit Problemen und Fehlern“ insbesondere bei der Beurteilung der Ergebnisse eines Experimentes relevant wird, wird dieser Modellaspekt im FSWEx.V2 der Subskala Auswertung zugeschrieben.

Tabelle 83: Operationalisierung der Teilkompetenzen Planung, Durchführung und Auswertung im Modell experimenteller Kompetenz (Schreiber et al. 2009) und im FSWEx.V2. Aufgeführt sind nur die Items, um die der FSWEx.V2 gegenüber dem FSWEx.V1 ergänzt wurde.

	Teilkompetenz/ Subskalen	Operationalisierung der Teilkompetenzen im Modell experimenteller Kompetenz nach Schreiber et al. 2009	Operationalisierung der Modell- Teilkompetenzen in den Subskalen des FSWEx.V2	Variablen- name
FSWEx.V2	Planung	Vorgegebene Fragestellungen klären & Versuchsplan entwerfen	Ich bin gut darin, mir zu überlegen, was für Geräte/Materialien ich für ein Experiment benötige.	pra9
		Vorgegebene Fragestellungen klären & Versuchsplan entwerfen	Es ist einfach für mich, mir eine Experimentieranleitung für die Beantwortung einer bestimmten Fragestellung zu überlegen.	pra15

	Planung	Vorgegebene Fragestellungen klären & Versuchsplan entwerfen	Wenn ich eine Idee für ein Experiment habe, fällt es mir leicht, mir einen geeigneten Versuchsaufbau dafür auszudenken.	pra29
	Durchführung	Versuchsanordnungen aufbauen	Der Aufbau von Versuchsanordnungen zum Experimentieren fällt mir leicht.	pra5
		Messungen durchführen	Ich bin gut darin, Experimente durchzuführen.	pra2
		Messungen durchführen	Es fällt mir leicht, die zu einem Experiment benötigten Geräte zu bedienen.	pra17
	Auswertung	Ergebnis interpretieren	Es ist einfach, Ergebnisse aus einem Experiment auszuwerten.	pra3
		Umgang mit Problemen und Fehlern	Bei einem nicht planmäßig verlaufenen Experiment fällt es mir leicht, mögliche Fehlerquellen zu finden.	pra24
		Messdaten aufbereiten	Es fällt mir leicht, aus experimentellen Messdaten Grafiken zu erstellen.	pra33

3.2.1.2.2 Prüfung der Fragebogengüte für den FSWEx.V2

Wie der Einsatz des erweiterten FSWEx.V2 im Pretest der Nebenstudie (inkl. Kontrollgruppe) zeigt, konnten die Reliabilitäten aller Teildimensionen auf Cronbachs $\alpha > 0,7$ gesteigert werden (Tab. 84).

Tabelle 84: Reliabilitäten der Teilkompetenzen Planung, Durchführung und Auswertung zum Zeitpunkt des Pretests im FSWEx.V2 der Nebenstudie (N ergibt sich aus den Schülern der Versuchs- und der Kontrollgruppe)

	Teilkompetenz	Anzahl Items	Reliabilität (Cronbachs α)	N
FSWEx.V2	Planung	6	0,789	177
	Durchführung	6	0,729	173
	Auswertung	6	0,786	177

Auch für den FSWEx.V2 wurde eine Faktorenanalyse durchgeführt, um zu prüfen, wie gut die Items die drei Subskalen des Modells nach Schreiber et al. (2009) repräsentieren.

3.2.1.2.3 Prüfung der Dimensionalität des FSWEx.V2

Bartlett-Test ($p \leq 0,001$) und ein Kaiser-Meyer-Olkin-Maß bestätigen ($KMO=0,887$) die Eignung der Variablen für die Durchführung einer Faktorenanalyse (Bühl 2010). Wie für den zuvor vorgestellten FSWEx.V1 wurde eine explorative Faktorenanalyse als Hauptkomponentenanalyse mit Varimax-

Rotation durchgeführt. Die Ladungen der einzelnen Items auf die entsprechenden FSWEx-Faktoren sind Tabelle 85 zu entnehmen.

Tabelle 85: Die in der explorativen Faktorenanalyse ermittelten Faktorladungen der FSWEx.V2-Items (Extraktionsmethode: Hauptkomponentenanalyse; Rotationsmethode: Varimax). Einheitlich farblich markiert sind die Items, die ihre höchste Ladung auf einem gemeinsamen Faktor aufweisen. Die rot markierten Items gruppieren anders als erwartet. Bei diesen Items werden auch schwache Faktorladungen aufgeführt (N=167).

		Komponenten		
Item		Faktor 1	Faktor 2	Faktor 3
Items zur Abbildung der Modelldimension Planung	pra11	0,482	0,289	0,053
	pra16	0,460	0,560	0,104
	pra27	0,708	0,115	0,176
	pra9	0,580	0,137	0,457
	pra15	0,651	0,252	0,238
	pra29	0,750	0,063	0,264
Items zur Abbildung der Modelldimension Planung	pra8	0,318	0,150	0,685
	pra22	-0,188	0,695	0,173
	pra31	-0,144	0,440	0,676
	pra2	0,164	0,172	0,682
	pra5	0,353	0,275	0,473
	pra17	0,218	-0,017	0,723
Items zur Abbildung der Modelldimension Durchführung	pra13	0,296	0,600	0,131
	pra18	0,275	0,647	0,269
	pra26	0,394	0,577	0,045
	pra3	0,483	0,475	0,109
	pra24	0,318	0,394	0,367
	pra33	0,192	0,616	0,140

Wie aus Tabelle 85 ersichtlich wird, bilden die Items zufriedenstellend die beabsichtigten Subskalen ab. Lediglich die Items pra16, pra22 und pra3 weisen die höchsten Ladungen auf anderen als den erwarteten Faktoren auf. Item pra22 und pra16 laden – wie bereits in der Faktorenanalyse für den FSWEx.V1 (vgl. Punkt 3.2.1.1.3) – deutlich mit den Items der Subskala Auswertung auf einen Faktor. Da eine Entfernung dieser Items im FSWEx.V2 aber keine deutliche Steigerung der Reliabilität erbringen würde, wurden sie in der finalen Auswertung zwecks besserer Modellabdeckung belassen. Item pra3 (Es ist einfach, Ergebnisse aus einem Experiment auszuwerten) lädt nur geringfügig stärker mit den Items der Subskala Planung auf einen Faktor als mit denen der Subskala Auswertung. Aus inhaltlichen Gründen verbleibt auch dieses Item für die weitere Auswertung in der beabsichtigten Subskala. In Orientierung an Mayer et al. (2008) werden im Anschluss wieder die drei faktoranalytisch begründeten Subskalen auf Interkorrelationen untersucht.

3.2.1.2.4 Interkorrelationen der Subskalen des FSWEx.V2

Moderate Interkorrelationen der Subskalen sprechen für eigenständige Unterdimensionen eines globaleren Konstruktes (Mayer et al. 2008). Die Interkorrelationen der drei Subskalen Planung, Durchführung und Auswertung in Tabelle 86 liegen zwischen 0,567 und 0,671 und damit im Bereich moderater Stärke. Somit unterstützen diese die Annahme, dass die FSWEx.V2-Subskalen Planung, Durchführung und Auswertung eigenständige Unterdimensionen des Konstruktes der experimentbezogenen Selbstwirksamkeitserwartung abbilden.

Tabelle 86: Interkorrelationen der FSWEx.V2-Subskalen Planung, Durchführung und Auswertung im Pretest (N=167)

	Planung	Durchführung	Auswertung
Planung	1	0,567	0,671
Durchführung	0,567	1	0,619
Auswertung	0,671	0,619	1

3.2.1.3 Entwicklung der Fragebogenversion postFSWEx.V1

Neben dem FSWEx.V1 und dem FSWEx.V2, welche die experimentbezogene Selbstwirksamkeitserwartung unabhängig von einer konkreten Experimentiersituation erfragen, wurde der postFSWEx.V1 entwickelt. Dieser erhebt – ebenfalls in drei Subskalen – die in einer konkreten Experimentiersituation wahrgenommene Experimentierkompetenz.

3.2.1.3.1 Theoriegeleitete Itemkonstruktion für den postFSWEx.V1

Der postFSWEx.V1 wurde ebenfalls auf Grundlage des Modells experimenteller Kompetenz nach Schreiber et al. (2009) entwickelt. Tabelle 87 gibt einen Überblick über die Zuordnung der Items zu den Modellaspekten.

Tabelle 87: Operationalisierung der Teilkompetenzen Planung, Durchführung und Auswertung im Modell experimenteller Kompetenz (Schreiber et al. 2009) und im postFSWEx.V1 (*invers kodiertes Item)

	Teilkompetenz/ Subskalen	Operationalisierung der Teilkompetenzen im Modell experimenteller Kompetenz nach Schreiber et al. 2009	Operationalisierung der Modell- Teilkompetenzen in den Subskalen des FSWEx.V2	Variablen- name
postFSWEx.V1	Planung	Fragestellungen entwickeln	Die durchgeführten Experimente haben bei mir Fragen aufgeworfen, die ich mit anderen Experimenten beantworten könnte.	poa46
		Hypothesen bilden	Ich hatte mir schon vor dem Experiment überlegt, was herauskommen müsste und lag mit meiner Überlegung (Hypothese) richtig.	poa38
		Versuchsplan entwerfen	Die Planung der durchgeführten Experimente wäre mir leicht gefallen.	poa29
	Durchführung	Messungen durchführen	Der Umgang mit den Laborgeräten war kein Problem.	poa40
		Messungen durchführen	Die Durchführung der Experimente ist mir leicht gefallen.	poa25
		Messungen dokumentieren	Mit dem Aufschreiben der Versuchsbeobachtungen hatte ich Schwierigkeiten.*	poa32*
	Auswertung	Messdaten aufbereiten/Messdaten verarbeiten.	Die Auswertung der Versuchsergebnisse war leicht.	poa35
		Ergebnis interpretieren	Es war kein Problem für mich, die Versuchsbeobachtungen zu deuten.	poa43
		Umgang mit Problemen und Fehlern	Ich hatte Schwierigkeiten bei der Interpretation der Versuchsergebnisse.*	poa17*

3.2.1.3.2 Prüfung der Fragebogengüte für den postFSWEx.V1

Der postFSWEx.V1 wurde bei allen Schülern eingesetzt, die im Rahmen der Hauptstudie das BeLL Bio besuchten. Tabelle 88 stellt die Reliabilitäten der Subskalen zusammen. Das Item poa32 (Mit dem Aufschreiben von Versuchsbeobachtungen hatte ich Schwierigkeiten) wurde in der finalen Auswertung der Hauptstudie aus der Subskala Durchführung entfernt. Die Reliabilität der Subskala konnte dadurch von Cronbachs $\alpha=0,526$ auf Cronbachs $\alpha=0,615$ gesteigert werden.

Tabelle 88: Reliabilitäten der Teilkompetenzen Planung, Durchführung und Auswertung zum Zeitpunkt des Posttests im postFSWEx.V1 der Hauptstudie

postFSWEx.	Teilkompetenz	Anzahl Items	Reliabilität (Cronbachs α)	N
postFSWEx.V1	Planung	3	0,614	474
	Durchführung	3(2)	0,526(0,615)	477(487)
	Auswertung	3	0,768	480

*In der finalen Datenauswertung wurde Item poa32 aus der Subskala Durchführung zur Reliabilitätsoptimierung entfernt. Die Reliabilität der Subskala steigt auf 0,615.

Wie für die beiden anderen FSWEx-Versionen, so wurde auch hier eine explorative Faktorenanalyse durchgeführt, um zu prüfen wie gut die Items drei eigenständige Subskalen repräsentieren.

3.2.1.3.3 Prüfung der Dimensionalität des postFSWEx.V1

Auch für dieses Messinstrument bestätigen Bartlett-Test ($p \leq 0,001$) und ein Kaiser-Meyer-Olkin-Maß (KMO=0,780) die Eignung der Variablen für die Durchführung einer Faktorenanalyse (vgl. Bühl 2010). Wie für den FSWEx.V1 und FSWEx.V2 wurde daher eine explorative Faktorenanalyse als Hauptkomponentenanalyse mit Varimax-Rotation durchgeführt. Die Ladungen der einzelnen Items auf die entsprechenden FSWEx-Faktoren sind Tabelle 89 zu entnehmen.

Tabelle 89: Die in der explorativen Faktorenanalyse ermittelten Faktorladungen der postFSWEx.V1-Items (Extraktionsmethode: Hauptkomponentenanalyse; Rotationsmethode: Varimax). Einheitlich farblich markiert sind die Items, die ihre höchste Ladung auf einem gemeinsamen Faktor aufweisen. Die rot markierten Items gruppieren anders als erwartet. Bei diesen Items werden auch schwache Faktorladungen aufgeführt (N=457).

		Komponenten		
		Faktor 1	Faktor 2	Faktor 3
Items zur Abbildung der Modelldimension Auswertung	poa29	0,205	0,612	0,215
	poa38	0,122	0,783	0,041
	poa46	-0,047	0,780	0,004
Items zur Abbildung der Modelldimension Planung	poa25	0,282	0,107	0,762
	poa32	0,677	-0,302	0,126
	poa40	0,127	0,068	0,860
Items zur Abbildung der Modelldimension Durchführung	poa17	0,775	0,135	0,111
	poa35	0,774	0,213	0,200
	poa43	0,718	0,240	0,201

Mit Ausnahme des Items poa32 (Mit dem Aufschreiben der Versuchsbeobachtungen hatte ich Schwierigkeiten) laden alle Items relativ eindeutig auf die erwarteten Faktoren. Item poa32 wurde, da es zu einer deutlichen Verschlechterung der Skalenreliabilität beitrug, für die weitere Auswertung aus der Subskala Durchführung entfernt. Das sehr ähnliche Items pra22 wurde aus identischem Grund aus der Auswertung des FSWEx.V1 gestrichen.

3.2.1.3.4 Interkorrelationen der Subskalen des postFSWEx.V1

Die Interkorrelationen der drei Subskalen Planung, Durchführung und Auswertung in Tabelle 90 liegen mit $r(469)=0,203$ bis $r(469)=0,446$ im mäßigen Bereich. Sie fallen etwas schwächer aus, als die Interkorrelationen im FSWEx.V1 und im FSWEx.V2. Somit wird hier noch deutlicher, dass die Skalen unterschiedliche Konstrukte abfragen (Mayer et al. 2008).

Tabelle 90: Interkorrelationen der postFSWEx.V1-Subskalen Planung, Durchführung und Auswertung im Posttest (N=469)

	Planung	Durchführung	Auswertung
Planung	1	0,203	0,302
Durchführung	0,203	1	0,446
Auswertung	0,302	0,446	1

3.2.1.4 Korrelationen der FSWEx-Subskalen mit dem biologischen Fähigkeitsselbstkonzept

Zur Beurteilung der Kriteriumsvalidität der FSWEx-Subskalen werden nachfolgend Korrelationen mit dem dispositionalen Fähigkeitsselbstkonzept hergestellt. Selbstwirksamkeit und Selbstkonzept stehen in engem Zusammenhang, so dass positive Korrelationen angenommen werden können. In die Berechnung der Korrelationen für die FSWEx.V1-Subskalen fließen alle an der Hauptstudie teilnehmenden Schüler zum Pretestzeitpunkt ein (Tabelle 91). Die Korrelationskoeffizienten liegen im mäßig hohen, positiven Bereich. Dies kann als Hinweis gewertet werden, dass die Subskalen des FSWEx.V1 und das Fähigkeitsselbstkonzept ein ähnliches, aber doch nicht identisches Konstrukt abbilden. Zwar kann das Außendriterium des Fähigkeitsselbstkonzeptes alleine die Validität nicht absichern, aber zumindest stützen.

Tabelle 91: Korrelationen der FSWEx.V1-Subskalen Planung, Durchführung und Auswertung mit dem Fähigkeitsselbstkonzept im Pretest der Versuchsgruppen der Hauptstudie (r =Pearson-Korrelation; N =Stichprobengröße)

	Planung	Fähigkeitsselbstkonzept
Planung	r N	0,389 607
Durchführung	r N	0,371 625
Auswertung	r N	0,493 615

Die Korrelierung der FSWEx.V2-Subskalen mit dem Fähigkeitsselbstkonzept kann nicht für die Schüler der Nebenstudien-Versuchsgruppe durchgeführt werden, da hier lediglich der FSWEx.V2 ohne weitere affektive Messinstrumente eingesetzt wurde. In die Korrelationen fließen daher nur die Daten von Schülern der Kontrollgruppe zum Zeitpunkt des Pretests ein. Eine Übersicht der Korrelationen liefert Tabelle 92. Die Höhe der Korrelationen liegt wie beim FSWEx.V2 im mäßigen Bereich, auch dies kann als Hinweis auf Kriteriumsvalidität gewertet werden.

Tabelle 92: Korrelationen der FSWEx.V2-Subskalen Planung, Durchführung und Auswertung mit dem biologischen Fähigkeitsselbstkonzept im Pretest der Kontrollgruppe (r =Pearson-Korrelation; N =Stichprobengröße)

	Planung	Fähigkeitsselbstkonzept
Planung	r N	0,336 128
Durchführung	r N	0,550 123
Auswertung	r N	0,554 126

Die Korrelationen für den postFSWEx.V1 wurden alle Schüler einbezogen, die im Rahmen der Hauptstudie das BeLL Bio besuchten. Da dieses Messinstrument im Posttest eingesetzt wurde, wird für die Berechnung entsprechend das im Posttest erhobene Fähigkeitsselbstkonzept berücksichtigt. Während die Korrelationen für die Subskalen Durchführung und Auswertung in ähnlicher Höhe liegen, wie bei den zuvor genannten FSWEx-Versionen, fällt die Korrelation in der Subskala relativ gering aus. Die dennoch positive Korrelation widerspricht nicht notwendigerweise der Validität, allerdings ist die niedrige Korrelation relativ unerwartet (Tab. 93).

Tabelle 93: Korrelationen der postFSWEx.V1-Subskalen Planung, Durchführung und Auswertung mit dem Fähigkeitsselbstkonzept im Posttest der Versuchsgruppen der Hauptstudie (r =Pearson-Korrelation; N =Stichprobengröße)

	Planung	Fähigkeitsselbstkonzept
Planung	r N	0,165 473
Durchführung	r N	0,361 497
Auswertung	r N	0,412 479

Der FSWEx wurde entwickelt, um die subjektiv wahrgenommene Experimentierkompetenz in einer Experimentiersituation zu erfassen. Ob die Schüler, welche sich fähiger einschätzen, in Experimentiersituationen tatsächlich kompetenter handeln, soll die Korrelation mit den Daten der Videoanalyse von Dahmen (2012) zeigen.

3.2.1.5 Korrelationen der FSWEx -Subskala Durchführung mit dem realen Experimentierverhalten

Um zu überprüfen, wie aus der Praxis erhobene Daten sich gegenüber diesen bisher theoretisch ermittelten Daten verhalten, wurden Filmsequenzen ausgewertet. Für den Abgleich der FSWEx-Daten mit der real im Schülerlabor beobachtbaren Experimentierfähigkeit standen Videodaten von einer Pipettierübung sowie vom Beladen eines Agarosegels mit DNA-Proben zur Verfügung. Diese Daten wurden während des Kurses Genetischer Fingerabdruck erhoben. Nachfolgend werden ausschließlich einseitig signifikante oder tendenzielle Korrelationen nach Kendall Tau zwischen der in Pre-, Post- und Follow-up-Test eingesetzten FSWEx-Subskala Durchführung und den Messkategorien der Videoanalyse vorgestellt. Einseitige Signifikanzen wurden deshalb gewählt, weil die gerichtete Annahme besteht, dass Schüler mit hoher experimentbezogener Selbstwirksamkeitserwartung auch entsprechend häufiger kompetentes Verhalten im Labor zeigen. Eine Korrelation des Laborverhaltens mit den FSWEx-Subskalen Planung und Auswertung wurde nicht vorgenommen, da die Videoaufnahmen ausschließlich die praktische Durchführung dokumentieren. Die Formulierung der Korrelationen in den folgenden Tabellen ergibt sich aus den Formulierungen im Auswerteschlüssel der Videoanalyse (vgl. Abschnitt 2.2.7.9).

Wie Tabelle 94 zu entnehmen ist, sind moderate Korrelationen zwischen der subjektiven Durchführungskompetenz in der Pretest-FSWEx.V1-Subskala Durchführung und der Häufigkeit von Bedienungsfehlern sowie der (Un-)sicherheit beim tatsächlichen Arbeiten im Labor festzustellen. Schüler, die ihre Experimentierkompetenz im Pretest höher einschätzen, machen in der Praxis weniger Fehler bei der Bedienung der Pipetten und sind mäßig sicher beim Beladen des Gels. Sie sind damit nicht als ganz sicher, aber auch nicht als unsicher einzustufen, was weder den Erwartungen entspricht, noch widerspricht.

Tabelle 94: Signifikante und tendenzielle Korrelationen (Kendall Tau) der im Pretest erhobenen FSWEx.V1-Subskala: Durchführung mit den Messdimensionen (V) der Videoanalyse (F: Messdimension Bedienungsfehler, U: Messdimension Unsicherheit im Arbeitsablauf; p =Signifikanz 1-seitig).

Pipettierübung					Gel beladen			
V	P	r	Verhalten im Labor		Verhalten im Labor	r	p	V
F	0,075	0,123	...machen weniger Bedienungsfehler.	Schüler, die im Pretest ihre Kompetenz, Experimente erfolgreich durchzuführen hoch einschätzen,...	...sind eher mäßig sicher beim Experimentieren.	0,148	0,055	U

Auch im postFSWEx.V1, in welchem die Schüler, rückblickend ihre Durchführungskompetenz im Laborkurs beurteilen, sind geringe Korrelationen zu beobachten. Diejenigen, die ihre

Durchführungskompetenz hoch einschätzen, sind im Laborkurs selten unsicher und handeln und kommunizieren häufiger Experiment-bezogen (Tabelle 95).

Tabelle 95: Signifikante und tendenzielle Korrelationen (Kendall Tau) der im Posttest erhobenen postFSWEx.V2-Subskala: Durchführung mit den Messdimensionen (V) der Videoanalyse (U: Messdimension Unsicherheit im Arbeitsablauf, K: Messdimension Kommunikation; p =Signifikanz 1-seitig).

Pipettierübung					Gel beladen			
V	P	r	Verhalten im Labor		Verhalten im Labor	r	p	V
U	0,073	0,131	...sind seltener unsicher beim Experimentieren.	Schüler, die im Posttest ihre Kompetenz, Experimente erfolgreich durchzuführen hoch einschätzen,...	...handeln häufiger Experiment-bezogen.	0,197	≤0,05	H
					...kommunizieren häufiger Experiment-bezogen.	0,110	0,078	K

Diejenigen Kursteilnehmer, die etwa zwei Monate nach dem Laborkurs ihre Experimentierkompetenz im FSWEx.V1 höher einschätzen, haben im Labor weniger Fehler bei der Bedienung der Pipetten gemacht und experimentbezogener gehandelt. Allerdings kommunizierten diese Schüler bei der Pipettierübung entgegen der Erwartungen tendenziell häufiger nicht Experiment-bezogen. Sowohl bei der Pipettierübung als auch beim Beladen des Gels war die Kommunikation der sich als kompetenter einschätzenden Schüler in der Regel zu beurteilen, während dies bei den sich weniger gut einschätzenden Kursteilnehmern nicht der Fall war. Die Beurteilbarkeit der Kommunikation hing nicht von der videographierten Person selbst ab, sondern wurde in der Regel durch Störgeräusche bedingt (Tab. 96).

Festzuhalten bleibt, dass die Korrelationen insgesamt nur gering ausfallen. Mit Ausnahme einer Korrelation im Follow-up-Test gehen diese aber konform mit den Selbsteinschätzungen im FSWEx.

Tabelle 96: Signifikante und tendenzielle Korrelationen (Kendall Tau) der im Follow-up-Test erhobenen FSWEx.V1-Subskala: Durchführung mit den Messdimensionen der Videoanalyse (V) (F: Messdimension Bedienungsfehler, K: Messdimension Kommunikation; p=Signifikanz 1-seitig).

Pipettierübung				Schüler, die im Follow-up-Test ihre Kompetenz, Experimente erfolgreich durchzuführen hoch einschätzen,...	Gel beladen			
V	P	r	Verhalten im Labor		Verhalten im Labor	r	p	V
F	≤0,05	0,173	...machen weniger Bedienungsfehler.		...handeln häufiger Experiment-bezogen.	0,132	0,089	H
K	0,060	0,153	...kommunizieren häufiger <u>nicht</u> Experiment-bezogen.		(Kommunikation meistens zu beurteilen) *	0,161	≤0,05	K
K	≤0,05	0,203	(Kommunikation meistens zu beurteilen) *					

*Die Kodierung *Kommunikation nicht zu beurteilen* wurde immer dann vorgenommen, wenn aufgrund schlechter Tonqualität oder Störgeräuschen im Video die Kommunikation nicht zu beurteilen war (Dahmen 2012).

3.2.2 Hauptstudie

Der Übersichtlichkeit halber erfolgt die Darstellung der in der Hauptstudie erhobenen Daten gegliedert in die Unterpunkte persönliche Disposition, Laborvariablen und Längsschnitteffekte. Im Abschnitt Längsschnitteffekte werden die über mehrere Messzeitpunkte erhobenen Daten präsentiert. Abschließend wird das evaluierte Gesamtmodell der affektiven Einflüsse am Lernort BeLL Bio vorgestellt. Differenziert wird zwischen den drei Laborkursen Genetischer Fingerabdruck, LDH und Einzeller dort, wo dies erforderlich erscheint. Da für den Kurs Einzeller keine Follow-up-Daten vorliegen, werden einzelne Auswertungen nur für die beiden zuerst genannten Laborkurse vorgenommen.

3.2.2.1 Persönliche Disposition

Unter der persönlichen Disposition werden hier sämtliche Variablen gefasst, welche die Schüler bereits vor dem Laborbesuch, also zum Zeitpunkt des Pretests kennzeichnen. Zu diesen zählen das Fach- und Sachinteresse, das bereits vor dem Laborbesuch bestehende biologische Fähigkeitsselbstkonzept sowie die im Pretest empfundene Selbstwirksamkeitserwartung, das Wissenschaftsverständnis, den empfundenen Beitrag von Experimenten zum eigenen Verständnis sowie das Vorwissen. Sämtliche Personenvariablen wurden im Pretest erhoben und werden in

diesem Abschnitt zunächst ohne Bezug auf die Einflüsse des Labors untersucht. Dies erscheint insbesondere in Hinblick auf die Identifizierung potentiell unterschiedlicher Lernertypen interessant. Die Daten der drei Treatmentgruppen Genetischer Fingerabdruck, LDH und Einzeller sowie die der Kontrollgruppe können daher in diesem Abschnitt zunächst gemeinsam betrachtet werden. Erst zum Ende dieses Abschnittes wird die Ausprägung der dispositionalen Variablen getrennt für die Treatmentgruppen und die Kontrollgruppe aufgezeigt. Neben den Ergebnissen, welche alle untersuchten Schüler einbeziehen, werden Subgruppenvergleiche anhand des Geschlechtes, des Migrationshintergrundes sowie der beruflichen Orientierung (naturwissenschaftlich vs. nicht naturwissenschaftlich) vorgenommen.

3.2.2.1.1 Fach-/Sachinteresse

Wie der folgende Abschnitt verdeutlichen wird, ist die gemeinsame Berücksichtigung des dispositionalen Interesses am Unterrichtsfach Biologie (Fachinteresse) und des Interesses am Experimentieren (Sachinteresse) von besonderer Relevanz für die Einschätzung der Heterogenität der Schüler. Diese Interessen werden in Abbildung 46 gegen die Biologienote – als Merkmal unterrichtlicher Leistungsfähigkeit – aufgetragen. Die dazu gehörigen Daten ergeben sich aus Tabelle 97 Die Schulnote wird hier als ordinalskaliert betrachtet, so dass als Korrelationskoeffizient Kendall Tau für die Korrelierung der Interessen mit der Biologienote verwendet wird. Was Abbildung 46 bereits augenscheinlich verrät, bestätigen die Korrelationen. Während das Fachinteresse höchst signifikant ($r(627)=-0,338$, $p\leq 0,001$) mit der Biologienote korreliert, ist ein Zusammenhang von Sachinteresse und der Biologienote nicht zu beobachten ($r(626)=0,001$, $p=0,976$). Während die schulisch als gut beurteilten Schüler also ein deutlich höheres Interesse am Biologieunterricht haben als die schlechter beurteilten, ist das Interesse am Experimentieren gleichermaßen als hoch einzustufen. Dieser Zusammenhang kann unabhängig vom Geschlecht, Migrationshintergrund oder beruflichen Plänen (naturwissenschaftlich/nicht naturwissenschaftlich) beobachtet werden.

Tabelle 97: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) der Schüler in den Skalen Fach- und Sachinteresse in Abhängigkeit von der Schulnote (vgl. Abbildung 46)

Bionote	Fachinteresse		Sachinteresse		N
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1	3,27	0,852	2,80	0,979	93
2	2,92	0,919	2,99	0,844	259
3	2,50	0,882	2,93	0,789	199
4	2,18	0,933	2,87	0,880	65
5	1,33	0,661	3,39	0,546	9
6	1,00	0,707	2,50	1,000	4

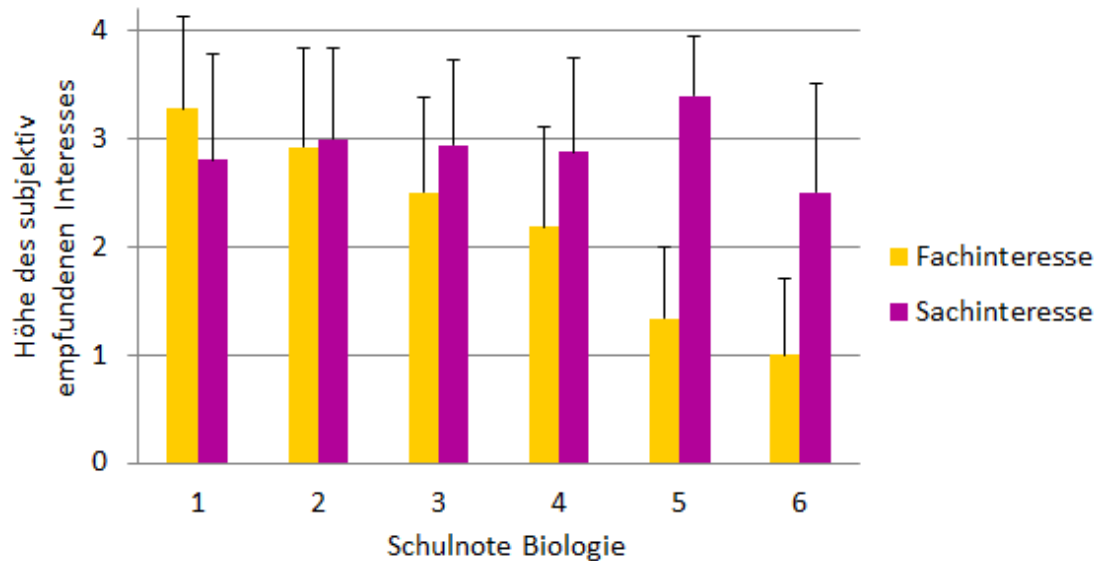


Abbildung 46: Abhängigkeit des Interesses am Unterrichtsfach Biologie (Fachinteresse) und am Experimentieren (Sachinteresse) von der Schulnote (Skalenmaximum=4)

Wie der folgende Vergleich des Fach- und Sachinteresses mit den anderen affektiven, dispositionalen Variablen aufzeigt, liegen alle Korrelationen – außer drei – im Bereich $r \leq 0,4$ (Tabelle 98). Höhere Korrelationen sind nur zwischen dem Fachinteresse und dem biologischen Fähigkeitsselbstkonzept ($r(628)=0,622$, $p \leq 0,001$) sowie zwischen dem Sachinteresse und der Variable Verständnis ($r(627)=0,553$, $p \leq 0,001$) und auch zwischen dem Sachinteresse und der FSWEx.V1-Subskala Durchführung von Experimenten ($r(625)=0,448$, $p \leq 0,001$) festzustellen. Die Selbsteinschätzung bezüglich der eigenen fachlichen Fähigkeit scheint demnach relativ eng mit dem Interesse am Biologieunterricht zusammenzuhängen. Das Sachinteresse scheint dagegen insbesondere mit der Überzeugung zu korrelieren, dass Experimente zum eigenen Verständnis beitragen sowie mit der Einschätzung, ein guter praktischer Experimentierer zu sein.

Tabelle 98: Pearson-Korrelationen (r), Signifikanzen (2-seitig) und Stichprobengröße (N) des Fach- und Sachinteresses mit den übrigen im Pretest erhobenen Personenvariablen. Korrelationen über $r \geq 0,4$ sind farblich gekennzeichnet.

		Selbstkonzept	FSWEx.V1 Planung	FSWEx.V1 Durchführung	FSWEx.V1 Auswertung	Verständnis	Wissenschafts- verständnis	Sachinteresse
Fach- interesse	r	0,622	0,382	0,261	0,361	0,368	0,301	0,270
	Signifikanz	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$
	N	628	607	625	616	628	629	628
Sach- interesse	r	0,103	0,260	0,448	0,120	0,553	0,253	1
	Signifikanz	$p \leq 0,01$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,01$	$p \leq 0,001$	$p \leq 0,001$	
	N	627	607	625	616	627	628	628

Dagegen scheinen das Wissenschaftsverständnis sowie die Selbstwirksamkeitserwartung Experimente erfolgreich planen und auswerten zu können, in keinem so deutlichen Zusammenhang zum Fach- und Sachinteresse zu stehen.

Die Beobachtungen treten unabhängig vom Geschlecht, Migrationshintergrund und beruflicher Orientierung in ähnlicher Form auf. Auffällig ist allerdings die relativ niedrige Korrelation zwischen Fähigkeitsselbstkonzept und Fachinteresse von Schülern, die planen in ein naturwissenschaftliches Berufsfeld einzusteigen ($r(104)=0,391$, $p\leq 0,001$) gegenüber den nicht naturwissenschaftlich orientierten Schülern ($r(268)=0,624$, $p\leq 0,001$). Die Höhe des Fach- und Sachinteresses unterscheidet sich zwischen männlichen und weiblichen Schülern nicht (Fachinteresse: $t(626)=1,10$; $p=0,270$; Sachinteresse: $t(626)=-0,02$; $p=0,984$). Auch eine Abhängigkeit vom Migrationshintergrund ist nicht festzustellen (Fachinteresse: $t(405)=0,61$; $p=0,544$; Sachinteresse: $t(405)=-0,68$; $p=0,494$). Während naturwissenschaftlich und nicht naturwissenschaftlich orientierte Schüler gleichermaßen Interesse am Experimentieren haben ($t(370)=1,80$; $p=0,073$), bekunden diejenigen mit naturwissenschaftlicher Ausrichtung ein deutlich höheres Fachinteresse ($t(370)=7,01$; $p\leq 0,001$; $\omega^2=0,115$, Tab. 99, Abb. 47).

Tabelle 99: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) im Fach- und Sachinteresse der Schüler in Abhängigkeit von der beruflichen Orientierung (vgl. Abbildung 47)

Subskala	Berufliche Orientierung	<i>M</i>	<i>SD</i>	<i>N</i>
Fachinteresse	naturwissenschaftlich	3,27	0,79	104
	nicht naturwissenschaftlich	2,57	0,91	268
Sachinteresse	naturwissenschaftlich	3,04	0,83	104
	nicht naturwissenschaftlich	2,87	0,86	268

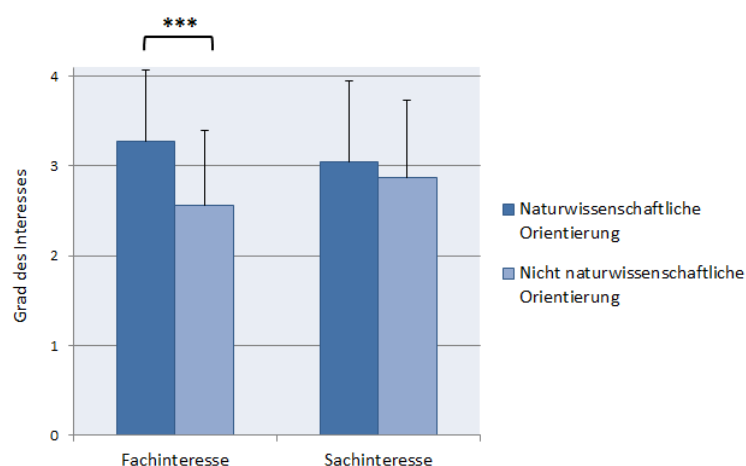


Abbildung 47: Mittelwerte und Standardabweichungen der Schüler im Fach- und Sachinteresse in Abhängigkeit von der beruflichen Orientierung (Skalenmaximum=4, *: $p\leq 0,001$)**

3.2.2.1.2 Biologisches Fähigkeitsselbstkonzept und experimentbezogene Selbstwirksamkeit (FSWEx.V1)

Im vorherigen Abschnitt wurde bereits der enge Zusammenhang des biologischen Fähigkeitsselbstkonzeptes mit dem Fachinteresse der Schüler aufgezeigt. Wie in Tabelle 100 dargestellt, ist weiterhin ein deutlicher Zusammenhang des fachlichen Selbstkonzeptes mit den Subskalen des FSWEx.V1 zu beobachten. Während dieser bei der Berücksichtigung aller Schüler insbesondere für die Subskala Auswertung deutlich wird ($r(615)=0,493$, $p\leq 0,001$), zeigt der Geschlechtervergleich bei den männlichen Schülern etwas höhere Korrelationen als bei den weiblichen Schülern. Zudem ist ein deutlicher Zusammenhang innerhalb der FSWEx.V1-Subskalen Planung und Auswertung ($r(604)=0,527$, $p\leq 0,001$) sowie zwischen der Subskala Planung und dem empfundenen Beitrag von Experimenten zum eigenen Verständnis zu belegen ($r(607)=0,450$, $p\leq 0,001$). Während die FSWEx.V1-Subskalen Planung und Auswertung nur gering mit dem Sachinteresse am Experimentieren korrelieren (Planung: $r(607)=0,260$, $p\leq 0,001$; Auswertung: $r(616)=0,120$, $p\leq 0,001$), ist der Zusammenhang zwischen der Subskala Durchführung und dem Sachinteresse mit $r(625)=0,448$, $p\leq 0,001$ deutlich höher. Schüler, die ihre Kompetenz höher einschätzen, Experimente erfolgreich durchzuführen, scheinen somit auch eher ein höheres Sachinteresse zu haben.

Tabelle 100: Pearson-Korrelationen (r), Signifikanzen (2-seitig) und Stichprobengröße (N) des biologischen Fähigkeitsselbstkonzeptes mit den übrigen im Pretest erhobenen Personenvariablen. Korrelationen über $r\geq 0,4$ sind farblich gekennzeichnet.

		Fachinteresse	FSWEx.V1 Planung	FSWEx.V1 Durchführung	FSWEx.V1 Auswertung	Verständnis	Wissenschafts- verständnis	Sachinteresse
Selbst- konzept	r Signifikanz N	0,622 $p\leq 0,001$ 628	0,389 $p\leq 0,001$ 607	0,371 $p\leq 0,001$ 625	0,493 $p\leq 0,001$ 615	0,297 $p\leq 0,001$ 627	0,329 $p\leq 0,001$ 628	0,103 $p\leq 0,001$ 628
FSWEx.V1 – Planung	r Signifikanz N	0,382 $p\leq 0,001$ 607	1	0,310 $p\leq 0,001$ 607	0,527 $p\leq 0,001$ 604	0,450 $p\leq 0,001$ 607	0,384 $p\leq 0,001$ 607	0,260 $p\leq 0,001$ 607
FSWEx.V1 – Durchfüh- rung	r Signifikanz N	0,261 $p\leq 0,001$ 625	0,261 $p\leq 0,001$ 625	1	0,398 $p\leq 0,001$ 615	0,371 $p\leq 0,001$ 625	0,247 $p\leq 0,001$ 625	0,448 $p\leq 0,001$ 625
FSWEx.V1 - Auswertung	r Signifikanz N	0,361 $p\leq 0,001$ 616	0,527 $p\leq 0,001$ 604	0,398 $p\leq 0,001$ 615	1	0,311 $p\leq 0,001$ 616	0,260 $p\leq 0,001$ 616	0,120 $p\leq 0,001$ 616

Die Höhe des biologischen Selbstkonzeptes zwischen den Geschlechtern ($t(542)=1,56$; $p=0,126$) sowie zwischen Schülern mit und ohne Migrationshintergrund ($t(405)=0,25$; $p=0,803$) unterscheidet sich im Pretest nicht. Schüler mit naturwissenschaftlicher beruflicher Orientierung weisen allerdings

ein höchst signifikant höheres Selbstkonzept auf als die nicht naturwissenschaftlich Orientierten ($t(267)=0,25$; $p\leq 0,001$; $\omega^2=0,152$). In den Subskalen des FSWEx.V1 sind keine vom Migrationshintergrund abhängigen Unterschiede festzustellen (Planung: $t(396)=0,16$; $p=0,877$; Durchführung: $t(396)=0,53$; $p=0,599$; Auswertung: $t(396)=0,46$; $p=0,650$). Beim Vergleich der Geschlechter sind signifikante Unterschiede lediglich in den Subskalen Planung ($t(503)=-3,73$; $p\leq 0,001$; $\omega^2=0,021$) und Auswertung ($t(602)=-2,77$; $p\leq 0,01$; $\omega^2=0,011$) zugunsten höherer Mittelwerte bei den männlichen Schülern zu beobachten. Ihre Kompetenz, Experimente erfolgreich durchführen zu können, schätzen weibliche und männliche Schüler dagegen gleich ein ($t(602)=-0,44$; $p=0,664$). In Anbetracht der niedrigen Effektstärke ist eine Abhängigkeit der experimentbezogenen Selbstwirksamkeit vom Geschlecht allerdings als eher gering einzustufen (Tab. 101, Abb. 48).

Tabelle 101: Mittelwerte (M), Standardabweichungen (SD) und Stichprobengrößen (N) der FSWEx.V1-Subskalen in Abhängigkeit vom Geschlecht der Schüler (vgl. Abbildung 48)

Subskala	Gruppe	M	SD	N
Planung	weiblich	1,65	0,64	348
	männlich	1,86	0,74	256
Durchführung	weiblich	2,70	0,76	348
	männlich	2,72	0,80	256
Auswertung	weiblich	2,25	0,67	348
	männlich	2,41	0,70	256

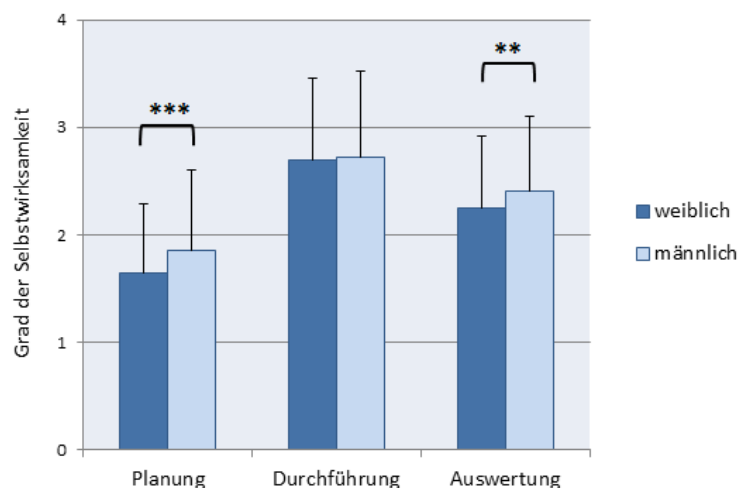


Abbildung 48: Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 für weibliche und männliche Schüler (Skalenmaximum=4, *: $p\leq 0,001$; **: $p\leq 0,01$)**

Die Schüler mit naturwissenschaftlicher beruflicher Orientierung weisen hingegen in allen drei Subskalen höchst signifikant höhere Werte auf als diejenigen, welche eine nicht naturwissenschaftliche Laufbahn anstreben (Planung: $t(362)=5,35$; $p\leq 0,001$; $\omega^2=0,070$; Durchführung: $t(362)=4,58$; $p\leq 0,001$; $\omega^2=0,052$; Auswertung: $t(396)=0,46$; $p\leq 0,001$; $\omega^2=0,098$). Damit übt die

Berufsorientierung einen stärkeren Effekt auf die experimentbezogene Selbstwirksamkeit aus als das Geschlecht (Tab. 102, Abb. 49).

Tabelle 102: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) der FSWEx.V1-Subskalen in Abhängigkeit von der beruflichen Orientierung der Schüler (vgl. Abbildung 49)

Subskala	Berufliche Orientierung	<i>M</i>	<i>SD</i>	<i>N</i>
Planung	naturwissenschaftlich	2,02	0,66	101
	nicht naturwissenschaftlich	1,62	0,63	263
Durchführung	naturwissenschaftlich	2,98	0,77	101
	nicht naturwissenschaftlich	2,57	0,74	263
Auswertung	naturwissenschaftlich	2,68	0,63	101
	nicht naturwissenschaftlich	2,20	0,64	263

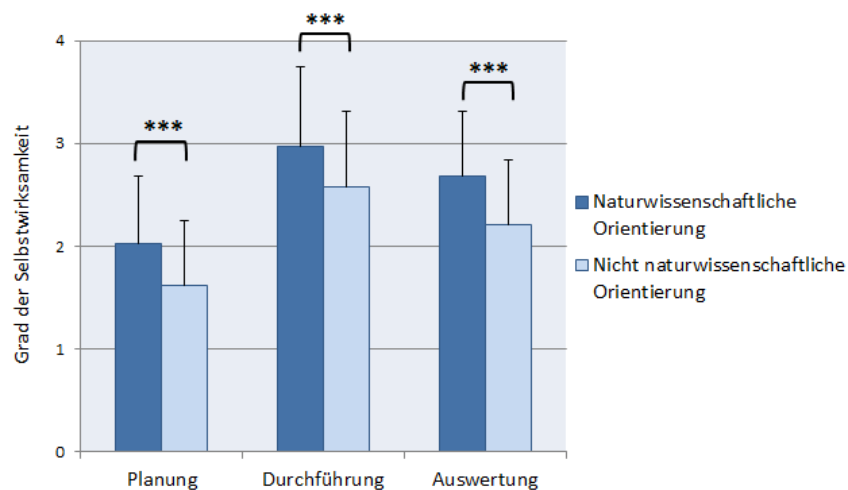


Abbildung 49: Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 für Schüler mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung (Skalenmaximum=4, *: $p \leq 0,001$).**

3.2.2.1.3 Verständnis und Wissenschaftsverständnis

Es wurde bereits festgestellt, dass das Sachinteresse am Experimentieren deutlich mit der Einschätzung korreliert, dass Experimente zum Verständnis beitragen können. Weiterhin können deutliche Korrelationen mit der FSWEx.V1-Subskala Planung ($r(607)=0,450$; $p \leq 0,001$) sowie dem Wissenschaftsverständnis ($r(628)=0,412$; $p \leq 0,001$) beobachtet werden (Tab. 103). Diese Korrelationen sind allerdings in den Subgruppen, die sich bei differenzierter Berücksichtigung des Geschlechtes, des Migrationshintergrundes oder der beruflichen Orientierung ergeben, nicht so stabil, wie die in allen Subgruppen bei $r \geq 0,50$ liegende Korrelation von Verständnis und Sachinteresse.

Tabelle 103: Pearson-Korrelationen (r), Signifikanzen (2-seitig) und Stichprobengröße (N) der Variable Verständnis mit den übrigen im Pretest erhobenen Personenvariablen. Korrelationen über $r \geq 0,4$ sind farblich gekennzeichnet.

		Fachinteresse	FSWEx.V1 Planung	FSWEx.V1 Durchführung	FSWEx.V1 Auswertung	Selbstkonzept	Wissenschafts- verständnis	Sachinteresse
Verständnis	r	0,368	0,450	0,371	0,311	0,297	0,412	0,553
	Signifikanz	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$
	N	628	607	625	616	627	628	627
Wissen- schafts- verständnis	r	0,301	0,384	0,247	0,260	0,247	1	0,253
	Signifikanz	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,001$		$p \leq 0,001$
	N	629	607	625	616	625		628

Zwischen männlichen und weiblichen Schülern (Verständnis: $t(626)=1,49$; $p=0,137$; Wissenschaftsverständnis: $t(626)=0,02$; $p=0,981$) sowie zwischen Schülern mit und ohne Migrationshintergrund (Verständnis: $t(404)=1,65$; $p=0,099$; Wissenschaftsverständnis: $t(404)=1,41$; $p=0,159$) ist weder im Wissenschaftsverständnis, noch im empfundenen Beitrag von Experimenten zum eigenen Verständnis ein statistisch bedeutsamer Unterschied festzustellen. Naturwissenschaftlich orientierte Schüler zeigen allerdings gegenüber nicht naturwissenschaftlich orientierten ein signifikant höheres Wissenschaftsverständnis ($t(369)=2,64$; $p \leq 0,01$; $\omega^2=0,016$). Außerdem messen die naturwissenschaftlich ausgerichteten Schüler dem Experiment einen signifikant höheren Beitrag zum Verständnis bei als die nicht naturwissenschaftlich orientierten ($t(369)=3,77$; $p \leq 0,001$; $\omega^2=0,034$). Wie die geringen Effektstärken zeigen, ist der Einfluss der Berufsorientierung auf die Variablen Verständnis und Wissenschaftsverständnis allerdings als eher gering einzustufen. Tabelle 104, Abbildung 50 und Abbildung 51 veranschaulichen die beschriebenen Unterschiede.

Tabelle 104: Mittelwerte (M), Standardabweichungen (SD) und Stichprobengrößen (N) der Schüler in den Variablen Verständnis (vgl. Abbildung 50) und Wissenschaftsverständnis (vgl. Abbildung 51) in Abhängigkeit von ihrer beruflichen Orientierung

Subskala	Berufliche Orientierung	M	SD	N
Verständnis	naturwissenschaftlich	2,74	0,80	104
	nicht naturwissenschaftlich	2,51	0,74	267
Wissenschaftsverständnis	naturwissenschaftlich	2,97	0,67	104
	nicht naturwissenschaftlich	2,70	0,58	267

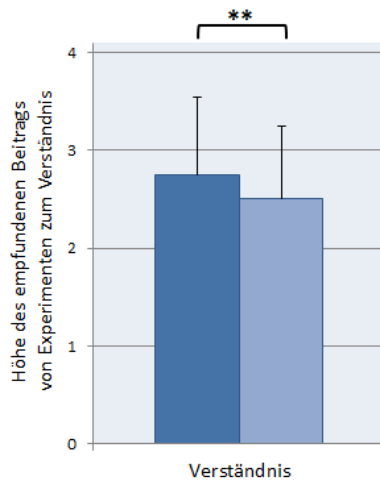


Abbildung 50: Durchschnittliche Höhe des empfundenen Beitrages von Experimenten zum Verständnis von Schülern mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung einschließlich der Standardabweichungen (Skalenmaximum=4, **: $p \leq 0,01$).

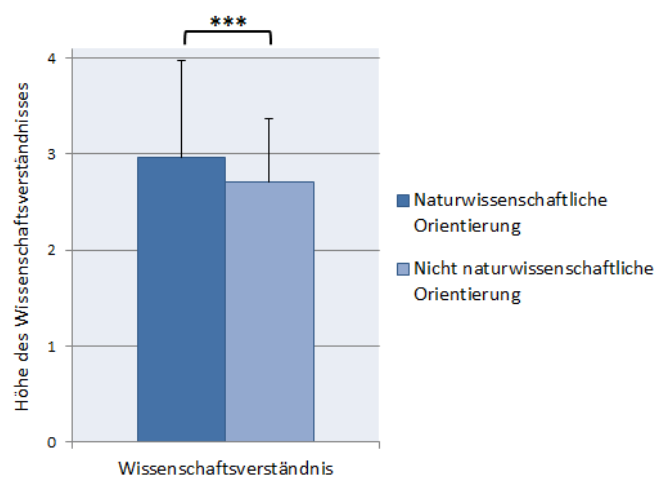


Abbildung 51: Mittelwerte und Standardabweichungen des Wissenschaftsverständnisses für Schüler mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung (Skalenmaximum=4, ***: $p \leq 0,001$).

3.2.2.1.4 Vorwissen

Neben den zuvor beschriebenen affektiven Variablen stellt auch das Vorwissen der Schüler eine dispositionale Personenvariable dar. Da sich die Wissenstests der drei Laborkurse inhaltlich unterscheiden, müssen die Korrelationen hier anhand des Wissenstests differenziert werden. Wieder werden zunächst Versuchs- und die entsprechende Kontrollgruppe gemeinsam betrachtet. Eine in Versuchs- und Kontrollgruppen differenzierte Betrachtung wird im Rahmen der Längsschnittdatenanalysen notwendig. Wie Tabelle 105 zeigt, korreliert das Wissen im Pretest Genetischer Fingerabdruck positiv, aber nur relativ moderat mit den affektiven dispositionalen Variablen.

Tabelle 105: Pearson-Korrelationen (r), Signifikanzen (2-seitig) und Stichprobengröße (N) der Testpunktezah im Wissenspretest Genetischer Fingerabdruck mit den affektiven dispositionalen Variablen.

		Selbstkonzept	Fachinteresse	Sachinteresse	Verständnis	Wissenschaftsverständnis	FSWEx.V1 Planung	FSWEx.V1 Durchführung	FSWEx.V1 Auswertung
Vorwissen Genet. Fingerabdruck	r	0,347	0,238	0,048	0,076	0,219	0,219	0,174	0,250
	Signifikanz	$p \leq 0,001$	$p \leq 0,001$	$p = 0,443$	$p = 0,225$	$p \leq 0,001$	$p \leq 0,001$	$p \leq 0,01$	$p \leq 0,001$
	N	255	255	255	255	255	249	255	252

Am höchsten korreliert das Vorwissen im Wissenspretest Genetischer Fingerabdruck mit dem Selbstkonzept ($r(255)=0,347$; $p \leq 0,001$). Schüler mit einem höheren Selbstkonzept erreichen in

diesem Test signifikant mehr Punkte als diejenigen, mit einem geringen Selbstkonzept. Die geringsten Korrelationen sind zwischen dem Wissenstest und dem Sachinteresse ($r(255)=0,048$; $p=0,048$) sowie der Variable Verständnis ($r(255)=0,076$; $p=0,225$) zu beobachten.

Aus Tabelle 106 geht hervor, dass die Testpunkte des Wissenstests LDH noch geringere Korrelationen zu den affektiven Variablen aufweisen als dies bei den Daten des Tests Genetischer Fingerabdruck der Fall war. Auch hier fallen die Korrelationen zwischen Wissen und Sachinteresse ($r(270)=0,010$; $p=0,833$) sowie Wissen und Verständnis ($r(270)=0,010$; $p=0,866$) wieder am geringsten aus. Allerdings sind auch die übrigen Korrelationen nicht bedeutend höher.

Tabelle 106: Pearson-Korrelationen (r), Signifikanzen (2-seitig) und Stichprobengröße (N) der Testpunktezah im Wissenspretest LDH mit den affektiven dispositionalen Variablen.

		Selbstkonzept	Fachinteresse	Sachinteresse	Verständnis	Wissenschafts- verständnis	FSWEx.V1 Planung	FSWEx.V1 Durchführung	FSWEx.V1 Auswertung
Vorwissen LDH	r	0,144	0,013	0,010	0,010	0,128	0,077	0,063	0,191
	Signifikanz	$p \leq 0,05$	$p=0,833$	$p=0,871$	$p=0,866$	$p \leq 0,05$	$p=0,218$	$p=0,304$	$p \leq 0,01$
	N	270	270	270	269	270	257	268	261

Noch geringer als in den ersten beiden Wissenstests fällt die Korrelation der im Wissenstest erreichten Punkte zu den affektiven dispositionalen Variablen in der Gruppe Einzeller aus (Tabelle 107). Hier ist keine der beobachteten Korrelationen signifikant.

Tabelle 107: Pearson-Korrelationen (r), Signifikanzen (2-seitig) und Stichprobengröße (N) der Testpunktezah im Wissenspretest Einzeller mit den affektiven dispositionalen Variablen.

		Selbstkonzept	Fachinteresse	Sachinteresse	Verständnis	Wissenschafts- verständnis	FSWEx.V1 Planung	FSWEx.V1 Durchführung	FSWEx.V1 Auswertung
Vorwissen Einzeller	r	0,090	0,008	0,117	0,028	-0,153	-0,096	0,123	-0,036
	Signifikanz	$p=0,366$	$p=0,933$	$p=0,238$	$p=0,781$	$p=0,121$	$p=0,339$	$p=0,220$	$p=0,720$
	N	103	104	103	104	104	101	102	103

Differenziert man die Korrelationsbetrachtungen nach Geschlecht und Migrationshintergrund⁷, so fallen die Korrelationen in beiden Subgruppen ähnlich gering aus. Bei den weiblichen Schülern, die den Wissenspretest Genetischer Fingerabdruck bearbeitet haben, korreliert allerdings die

⁷ Diese Differenzierung ist nicht für die Schüler möglich, welche den Wissenstest Einzeller ausgefüllt haben, da Migrationshintergrund und berufliche Orientierung erst im Follow-up-Test erfragt wurden. Dieser wurde für die Gruppe Einzeller nicht durchgeführt.

Testpunktezahl höchst signifikant mit ihrem Fachinteresse ($r(143)=0,36$; $p\leq 0,001$). Bei den männlichen Schülern ist hier kein bedeutender Zusammenhang zu beobachten ($r(112)=0,08$; $p=0,393$).

Bedeutsame Unterschiede in der Höhe des Vorwissens für Schüler mit und ohne Migrationshintergrund sind in den drei Wissenstests nicht zu beobachten. Auch beim Vergleich der Geschlechter sind die Unterschiede gering. Lediglich die weiblichen Schüler, die den Wissenstest LDH bearbeiteten, schneiden geringfügig schlechter ab als die männlichen Schüler ($t(195)=-2,68$; $p\leq 0,01$; $\omega^2=0,031$, Tab. 108, Abb. 52 bis 54). Die geringe Effektstärke deutet allerdings nur einen geringen Einfluss des Geschlechtes auf das Vorwissen an.

Tabelle 108: Mittelwerte (M), Standardabweichungen (SD) und Stichprobengrößen (N) der Schüler im Wissenspretest Genetischer Fingerabdruck, LDH und Einzeller, differenziert nach Geschlecht. Berücksichtigt sind Schüler der Versuchs- und Kontrollgruppen, welche zu allen drei Messzeitpunkten anwesend waren (vgl. Abbildungen 52 bis 54).

	Geschlecht	M	SD	N
Vorwissen Gen. Finger	weiblich	31,61	7,27	111
	männlich	32,16	5,96	88
Vorwissen LDH	weiblich	24,27	4,11	118
	männlich	25,87	4,10	79
Vorwissen Einzeller	weiblich	30,80	4,53	60
	männlich	29,98	3,88	46

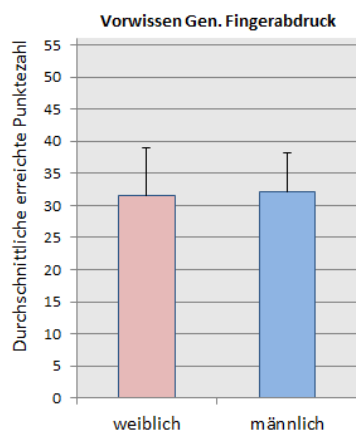


Abbildung 52: Vorwissen im Wissenspretest Genetischer Fingerabdruck differenziert nach Geschlecht inkl. Standardabweichung (Skalenmaximum=56 Punkte)

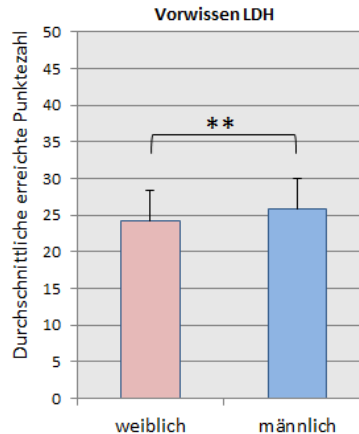


Abbildung 53: Vorwissen im Wissenspretest LDH differenziert nach Geschlecht inkl. Standardabweichung (Skalenmaximum=50 Punkte, ** $p\leq 0,01$)

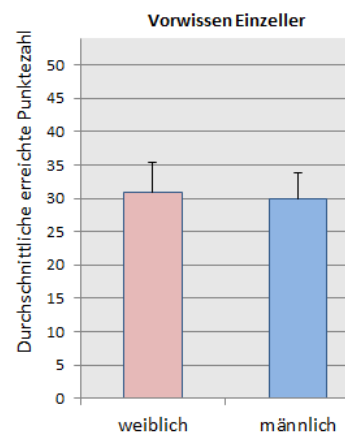


Abbildung 54: Vorwissen im Wissenspretest Einzeller differenziert nach Geschlecht inkl. Standardabweichung (Skalenmaximum=54 Punkte)

Etwas deutlichere Unterschiede im Vorwissen sind zwischen Schülern in Abhängigkeit von ihrer beruflichen Orientierung zu finden. Naturwissenschaftlich ausgerichtete Schüler schneiden sowohl im Wissenspretest Genetischer Fingerabdruck ($t(180)=4,71$; $p\leq 0,001$; $\omega^2=0,104$, als auch LDH-Test

($t(176)=2,99$; $p\leq 0,01$; $\omega^2=0,043$) besser ab als ihre nicht naturwissenschaftlich orientierten Mitschüler (Tab. 109, Abb. 55 bis 56).

Tabelle 109: Mittelwerte (M), Standardabweichungen (SD) und Stichprobengrößen (N) der Schüler im Wissenspretest Genetischer Fingerabdruck, LDH und Einzeller, differenziert nach beruflicher Orientierung. Berücksichtigt sind Schüler der Versuchs- und Kontrollgruppen, welche zu allen drei Messzeitpunkten anwesend waren (vgl. Abbildungen 55 und 56)

	Geschlecht	M	SD	N
Vorwissen Gen. Finger	naturwissenschaftlich	35,46	6,80	50
	nicht naturwissenschaftlich	30,43	6,29	132
Vorwissen LDH	naturwissenschaftlich	26,38	4,79	48
	nicht naturwissenschaftlich	24,28	3,90	130

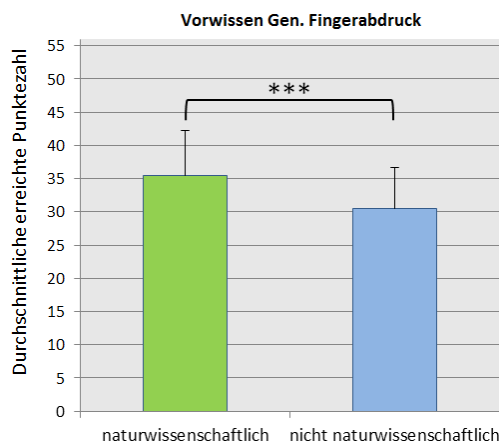


Abbildung 55: Vorwissen im Wissenspretest Genetischer Fingerabdruck differenziert nach beruflicher Orientierung inkl. Standardabweichung (Skalenmaximum=56 Punkte)

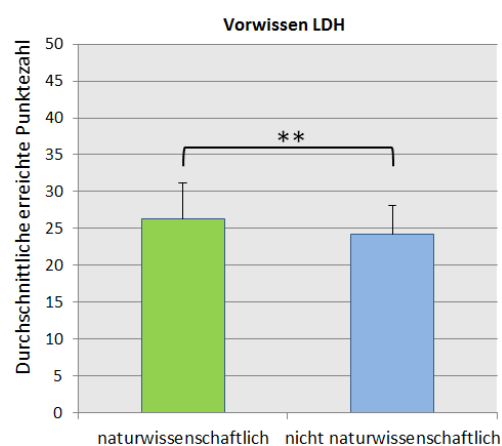


Abbildung 56: Vorwissen im Wissenspretest LDH differenziert nach beruflicher Orientierung inkl. Standardabweichung (Skalenmaximum=50 Punkte)

Eine Aussage über den Einfluss des Migrationshintergrundes und die berufliche Orientierung kann für den Wissenstest Einzeller nicht getroffen werden, da entsprechende Informationen hier nicht erhoben wurden. Differenzierter wird das Vorwissen im Vergleich der Versuchsgruppen mit der Kontrollgruppe sowie im Rahmen der Längsschnittanalysen unter Punkt 3.2.2.3.1. untersucht.

3.2.2.1.5 Schülertypen

Im vorangegangenen Text wurden bivariate Korrelationen verwendet, um den gegenseitigen Einfluss der Personenvariablen zu beschreiben. Eine hierarchische Clusteranalyse unter Verwendung der Clustermethode *Complete Linkage* (Entferntester Nachbar) soll nun zeigen, ob sich anhand dieser Variablen Schülertypen identifizieren lassen, die bestimmte Variablenausprägungen vereinen. In die Analyse gehen alle im Pretest erhobenen affektiven Personenvariablen ein. An dieser Stelle wird auf die Berücksichtigung der Wissensdaten verzichtet, da diese laborkursspezifisch erhoben wurden. Die identifizierten Schülertypen sollen aber kursübergreifend und auch für die Kontrollgruppe Geltung

besitzen. Im Rahmen der Clusteranalyse konnten drei Schülertypen identifiziert werden, die im Weiteren als Allrounder, Praktiker und Theoretiker bezeichnet werden. Tabelle 110 und Abbildung 57 stellen die Mittelwerte der Personenvariablen der drei Cluster zusammen. In der Abbildung sind die Mittelwerte jeder Variablen durch Linien verbunden, um Unterschiede zwischen den Clustern zu veranschaulichen.

Tabelle 110: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) der Pretest-Personenvariablen der drei Cluster Allrounder, Praktiker, Theoretiker (vgl. Abbildung 57)

Skala	Cluster	<i>M</i>	<i>SD</i>	<i>N</i>
Selbstkonzept	Allrounder	3,28	0,69	306
	Praktiker	1,77	0,96	153
	Theoretiker	2,93	0,66	145
Fachinteresse	Allrounder	3,36	0,57	306
	Praktiker	1,77	0,74	153
	Theoretiker	2,50	0,83	145
Sachinteresse	Allrounder	3,30	0,63	306
	Praktiker	3,06	0,72	153
	Theoretiker	2,13	0,76	145
Verständnis	Allrounder	2,92	0,63	306
	Praktiker	2,41	0,78	153
	Theoretiker	2,08	0,64	145
Wissenschaftsverständnis	Allrounder	3,01	0,60	306
	Praktiker	2,56	0,62	153
	Theoretiker	2,57	0,53	145
FSWEx.V1 - Planung	Allrounder	2,09	0,64	306
	Praktiker	1,40	0,60	153
	Theoretiker	1,36	0,49	145
FSWEx.V1 - Durchführung	Allrounder	2,97	0,70	306
	Praktiker	2,60	0,82	153
	Theoretiker	2,28	0,65	145
FSWEx.V1 - Auswertung	Allrounder	2,55	0,68	306
	Praktiker	1,92	0,66	153
	Theoretiker	2,23	0,49	145

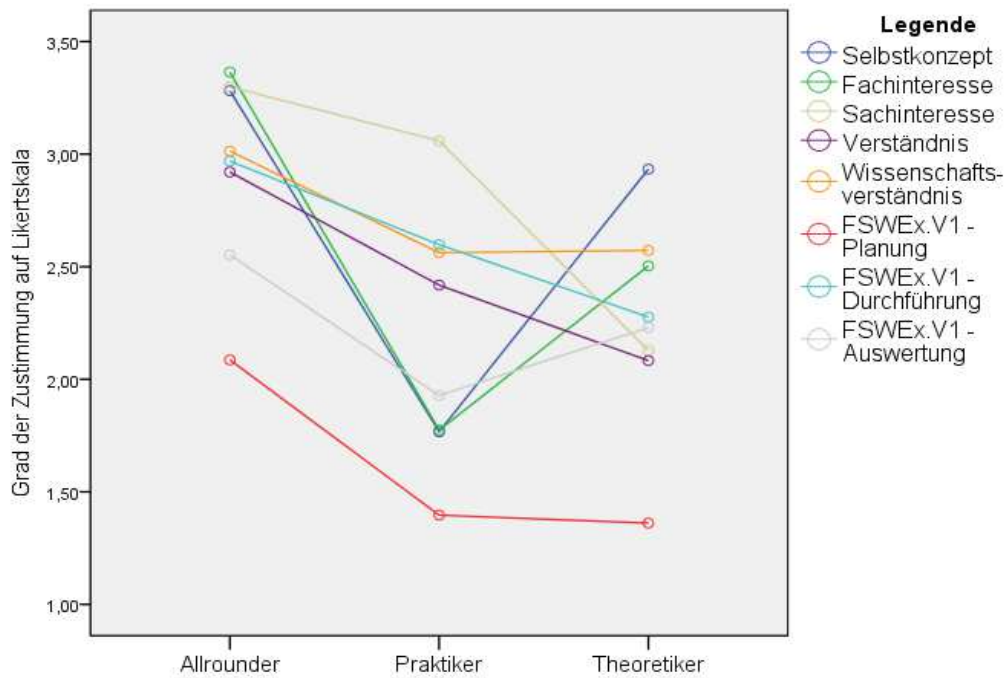


Abbildung 57: Mittelwerte in den affektiven dispositionalen Personenvariablen der drei aus der Clusteranalyse hervorgegangenen Gruppen Allrounder, Praktiker und Theoretiker. Zur Veranschaulichung von Unterschieden der drei Cluster wurden die Mittelwerte durch Linien verbunden (Skalenmaximum=4).

Die Gruppe der Allrounder setzt sich aus Schülern zusammen, die im Vergleich zu den anderen beiden Clustern in allen Personenvariablen die höchsten Ausprägungen besitzen. Besonders auffällig sind ihr hohes fachliches Interesse sowie ihr ausgeprägtes fachliches Selbstkonzept bei gleichzeitig hohem Sachinteresse am Experimentieren. Weiterhin schätzen diese Schüler ihre experimentbezogene Selbstwirksamkeit hoch ein. Sie zeigen deutliches Wissenschaftsverständnis, und das Experimentieren trägt für sie deutlich zum Verständnis bei. Mit einer Durchschnittsnote von 2,2 weisen die Allrounder den besten Notenschnitt der drei Schülertypen im Unterrichtsfach Biologie auf.

Das Cluster der Praktiker hebt sich deutlich von der Gruppe der Allrounder ab. Die Schüler dieses Clusters zeigen zwar ein ähnlich hohes Sachinteresse, ihr Fachinteresse sowie ihr fachliches Selbstkonzept stufen sie aber besonders gering ein. Interessant ist, dass die Praktiker ihre Selbstwirksamkeit, Experimente durchzuführen – im Vergleich zu den Theoretikern – relativ hoch, ihre Selbstwirksamkeit, Experimente zu planen und auszuwerten dagegen sehr niedrig einschätzen. Für die Praktiker trägt das Experiment zudem deutlicher zum Verständnis bei als für die Theoretiker, während ihr Wissenschaftsverständnis auf ähnlichem Niveau liegt. Die durchschnittliche Biologienote der Praktiker liegt mit 3,0 deutlich unter der der Allrounder.

Das Fachinteresse sowie das fachliche Selbstkonzept der Theoretiker liegen deutlich über dem der Praktiker, wenn auch nicht auf ganz so hohem Niveau wie dies bei den Allroundern der Fall ist. Die Theoretiker stufen ihr Sachinteresse am Experimentieren mit Abstand am niedrigsten ein und auch

ihre Selbstwirksamkeitserwartung, Experimente erfolgreich durchzuführen, ist verglichen mit den Allroundern und Praktikern gering. Für die Theoretiker tragen Experimente wenig zum Verständnis bei, ihr Wissenschaftsverständnis liegt aber auf ähnlichem Niveau wie das der Praktiker. Während die Selbstwirksamkeitserwartung für die Planung von Experimenten ähnlich hoch ist wie die der Praktiker, schätzen die Theoretiker ihre Selbstwirksamkeit höher ein, Experimente erfolgreich auswerten zu können. Mit einer durchschnittlichen Biologienote von 2,4 liegt die schulisch eingeschätzte Leistung der Theoretiker im Unterrichtsfach Biologie zwischen der der Praktiker und der Allrounder.

Mittels einfaktorieller ANOVA wurden die dispositionalen Personenvariablen auf statistisch bedeutsame Unterschiede zwischen den drei Clustern untersucht. Diese belegt höchst signifikante Unterschiede zwischen den Gruppen für jede Personenvariable (FSWEx.V1-Planung: $F(2,601)=107,00$; $p \leq 0,001$; $\eta^2=0,263$; FSWEx.V1-Durchführung: $F(2,601)=47,94$; $p \leq 0,001$; $\eta^2=0,138$; FSWEx.V1-Auswertung: $F(2,601)=51,42$; $p \leq 0,001$; $\eta^2=0,146$; Selbstkonzept: $F(2,601)=205,52$; $p \leq 0,001$; $\eta^2=0,406$; Fachinteresse: $F(2,601)=285,74$; $p \leq 0,001$; $\eta^2=0,487$; Sachinteresse: $F(2,601)=145,23$; $p \leq 0,001$; $\eta^2=0,326$; Verständnis: $F(2,601)=83,75$; $p \leq 0,001$; $\eta^2=0,218$; Wissenschaftsverständnis: $F(2,601)=43,63$; $p \leq 0,001$; $\eta^2=0,127$). Um festzustellen, welches Cluster sich von welchem unterscheidet, wurden für alle dispositionalen Personenvariablen Post-hoc-Tests nach Scheffé durchgeführt. Diese bestätigten für jede Variable mindestens sehr signifikante (Scheffé; $p \leq 0,01$) Unterschiede im Kreuzvergleich der drei Cluster mit Ausnahme von nicht signifikanten Ergebnissen zwischen den Theoretikern und Praktikern in der FSWEx.V1-Subskala Planung (Scheffé; $p=0,875$) sowie im Wissenschaftsverständnis (Scheffé; $p=0,989$). Werden die Ausprägungen der Personenvariablen für die drei Treatmentgruppen Genetischer Fingerabdruck, LDH und Einzeller sowie die Kontrollgruppe getrennt betrachtet, so bleiben die zuvor beschriebenen Verhältnisse weitgehend erhalten (Abbildung 58 A bis D). Die Daten zur Abbildung befinden sich in Anhang B.1.

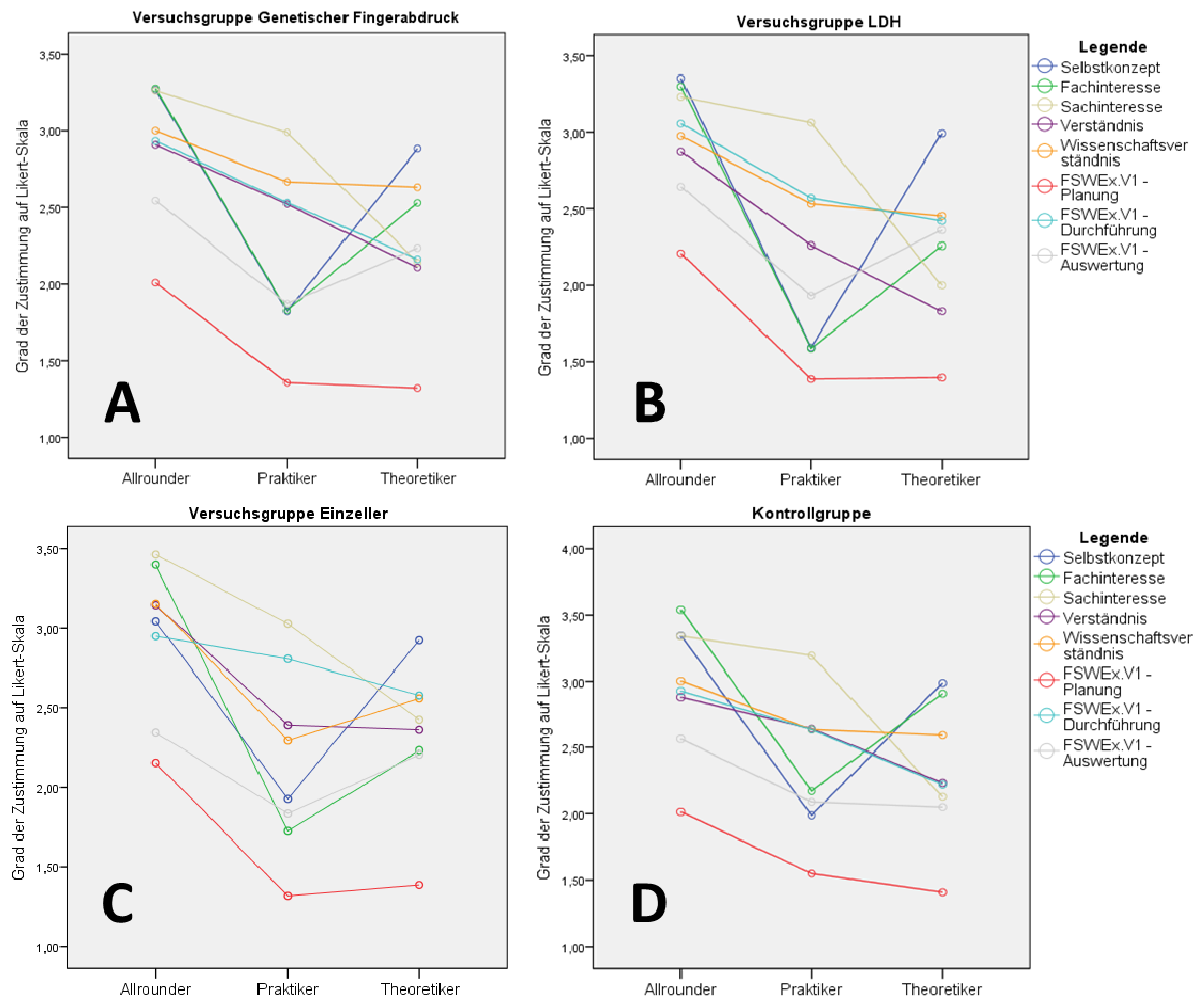


Abbildung 58: A-D: Mittelwerte in den affektiven dispositionalen Personenvariablen der drei aus der Clusteranalyse hervorgegangenen Gruppen Allrounder, Praktiker und Theoretiker für die Versuchsgруппe Genetischer Fingerabdruck (A), die Versuchsgруппe LDH (B), die Versuchsgруппe Einzeller (C) und die Kontrollgruppe (D) (Skalenmaximum=4).

Sowohl in den Versuchsgруппen, als auch in der Kontrollgruppe weisen die Allrounder und die Theoretiker ein deutlich höheres biologisches Fähigkeitsselbstkonzept und ein deutlich höheres Fachinteresse auf, als die Praktiker. Die Praktiker fallen in allen Gruppen durch ein gegenüber den Theoretikern deutlich höheres Interesse am Experimentieren auf.

Wie Tabelle 111 zeigt, sind die drei Cluster in den Versuchsgруппen sowie in der Kontrollgruppe ausreichend repräsentiert. Um eventuelle Unterschiede in der Wirkung der drei Laborkurse herausarbeiten zu können, werden im Weiteren neben dem Geschlecht, dem Migrationshintergrund und der beruflichen Orientierung auch differenziertere Vergleiche der drei Cluster durchgeführt.

Tabelle 111: Anzahl und prozentualer Anteil der Allrounder, Praktiker und Theoretiker in den drei Versuchsgruppen und in der Kontrollgruppe

		Allrounder	Praktiker	Theoretiker	Gesamt
Versuchsgruppe Gen. Fingerabdruck	Anzahl	104	46	68	218
	%	47,7%	21,1%	31,2%	100,0%
Versuchsgruppe Einzeller	Anzahl	39	18	13	70
	%	55,7%	25,7%	18,6%	100,0%
Versuchsgruppe LDH	Anzahl	88	63	37	188
	%	46,8%	33,5%	19,7%	100,0%
Kontrollgruppe	Anzahl	75	26	27	128
	%	58,6%	20,3%	21,1%	100,0%
Gesamt	Anzahl	306	153	145	604
	%	50,7%	25,3%	24,0%	100,0%

3.2.2.1.6 Vergleich von Versuchsgruppen und Kontrollgruppe

Für die weitere Datenauswertung ist es sinnvoll festzustellen, in welchen Variablen sich die Versuchsgruppen bereits im Pretest deutlich von der Kontrollgruppe unterscheiden. Tabelle 112 und Abbildung 59 geben zunächst einen Überblick über die im Pretest erzielten Mittelwerte der drei Versuchsgruppen sowie der Kontrollgruppe.

Tabelle 112: Mittelwerte (M) und Standardabweichungen (SD) der affektiven dispositionalen Variablen im Pretest für die drei Laborkurse und die Kontrollgruppe (vgl. Abbildung 59).

	Treatmentgruppe	M	SD	N
Selbstkonzept	Versuchsgruppe Gen. Fingerprint	2,84	0,92	218
	Versuchsgruppe Einzeller	2,73	0,99	70
	Versuchsgruppe LDH	2,69	1,11	188
	Kontrollgruppe	2,99	0,87	128
Fachinteresse	Versuchsgruppe Gen. Fingerprint	2,74	0,86	218
	Versuchsgruppe Einzeller	2,75	1,00	70
	Versuchsgruppe LDH	2,52	1,07	188
	Kontrollgruppe	3,13	0,79	128
Sachinteresse	Versuchsgruppe Gen. Fingerprint	2,86	0,84	218
	Versuchsgruppe Einzeller	3,16	0,74	70
	Versuchsgruppe LDH	2,93	0,88	188
	Kontrollgruppe	3,05	0,78	128
Verständnis	Versuchsgruppe Gen. Fingerprint	2,58	0,72	218
	Versuchsgruppe Einzeller	2,80	0,66	70
	Versuchsgruppe LDH	2,46	0,84	188
	Kontrollgruppe	2,70	0,71	128
Wissenschaftsverständnis	Versuchsgruppe Gen. Fingerprint	2,82	0,59	218
	Versuchsgruppe Einzeller	2,82	0,65	70
	Versuchsgruppe LDH	2,72	0,66	188
	Kontrollgruppe	2,84	0,62	128

FSWEx.V1 - Planung	Versuchsgruppe Gen. Fingerprint	1,66	0,65	218
	Versuchsgruppe Einzeller	1,80	0,72	70
	Versuchsgruppe LDH	1,77	0,73	188
	Kontrollgruppe	1,79	0,67	128
FSWEx.V1 - Durchführung	Versuchsgruppe Gen. Fingerprint	2,61	0,80	218
	Versuchsgruppe Einzeller	2,84	0,68	70
	Versuchsgruppe LDH	2,77	0,81	188
	Kontrollgruppe	2,72	0,72	128
FSWEx.V1 - Auswertung	Versuchsgruppe Gen. Fingerprint	2,31	0,68	218
	Versuchsgruppe Einzeller	2,19	0,62	70
	Versuchsgruppe LDH	2,35	0,74	188
	Kontrollgruppe	2,36	0,64	128

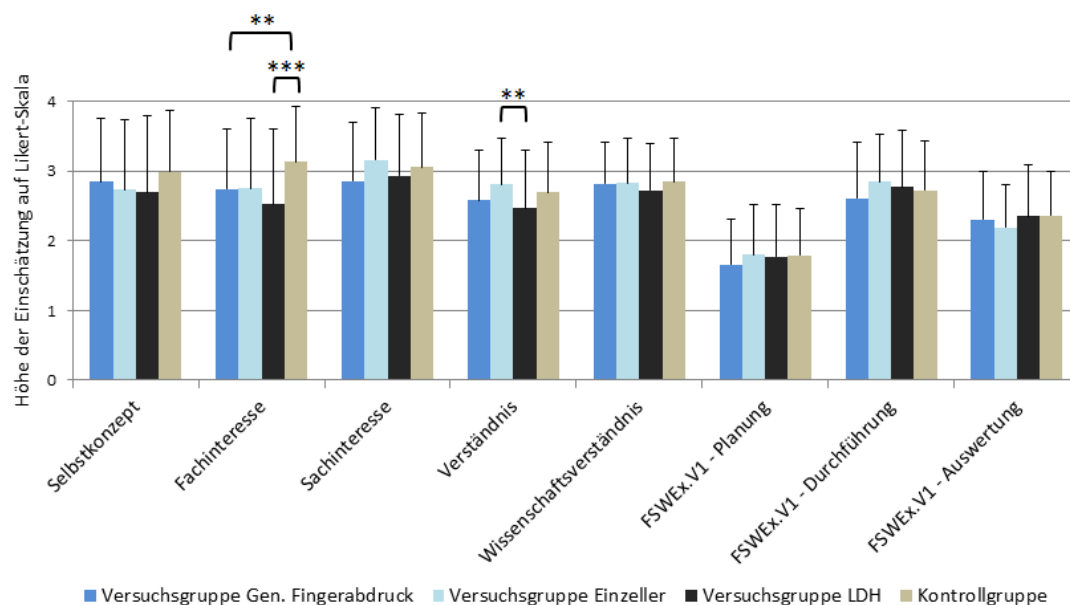


Abbildung 59: Durchschnittliche Höhe der Einschätzungen der Schüler zur Wahrnehmung der affektiven dispositionalen Variablen im Pretest in den drei Laborkursen sowie in der schulischen Kontrollgruppe inklusive Standardabweichungen (Skalenmaximum=4)

Wie Abbildung 59 zeigt, scheint die Ausprägung der dispositionalen Variablen in den untersuchten Gruppen insgesamt relativ homogen zu sein. Lediglich im Fachinteresse zeichnet sich ab, dass die Schüler der Kontrollgruppe hier gegenüber den Versuchsgruppen deutlich höhere Einschätzungen abgaben. Einfaktorielle ANOVAs bestätigten signifikante Unterschiede zwischen den Vergleichsgruppen im Fachinteresse ($F(3,600)=10,74$; $p \leq 0,001$; $\eta^2=0,051$), im Sachinteresse ($F(3,600)=3,08$; $p \leq 0,05$; $\eta^2=0,015$) sowie im empfundenen Beitrag von Experimenten zum Verständnis ($F(3,600)=4,63$; $p \leq 0,01$; $\eta^2=0,023$). Zur Feststellung, zwischen welchen Gruppen die beobachteten Unterschiede auftreten, wurden post-hoc-Tests nach Scheffé durchgeführt. Diese bestätigten einen signifikanten Unterschied im Fachinteresse zwischen der Kontrollgruppe und der Versuchsgruppe Genetischer Fingerabdruck (Scheffé; $p \leq 0,01$) sowie der Kontrollgruppe und der Versuchsgruppe LDH (Scheffé; $p \leq 0,001$). Die Kontrollgruppe unterscheidet sich im Fachinteresse nur

tendenziell von der Gruppe Einzeller (Scheffé; $p=0,061$). Im Sachinteresse können mittels Scheffé-Test keine signifikanten Unterschiede zwischen den Gruppen festgestellt werden. Lediglich die Versuchsgruppe Einzeller unterscheidet sich hier tendenziell von der Gruppe Genetischer Fingerabdruck (Scheffé; $p=0,074$). In der Variable Verständnis liegt der beobachtete Unterschied zwischen den Treatmentgruppen Einzeller und LDH (Scheffé; $p\leq 0,01$). Weitere statistisch bedeutsame Unterschiede zwischen den Vergleichsgruppen sind nicht festzustellen. Selbst für das augenscheinlich auffällige Fachinteresse liegt der beobachtete Unterschied im Bereich geringer Effektstärke. Da ein Einfluss des vor dem Laborkurs bestehenden Fachinteresses auf das aktuelle Interesse nach dem Laborbesuch angenommen wird, wird das Fachinteresse – trotz der nur geringen beobachteten Unterschiede – in der entsprechenden Längsschnittanalyse als Kovariate berücksichtigt.

Zwar fallen die Unterschiede der affektiven Variablen zwischen den Vergleichsgruppen gering aus, zu prüfen bleibt aber, ob dies auch für die kognitive Dimension gilt. Einen Überblick über die im Pretest erreichte Punktezahl im Wissenstest für die drei Versuchsgruppen sowie für die der entsprechenden Kontrollgruppe findet sich in Tabelle 113 sowie in den Abbildungen 60 bis 62. Da dies für die Längsschnittanalysen relevant wird, fließen in die Auswertung ausschließlich die Schüler ein, welche auch zu den späteren Messzeitpunkten anwesend sind. Jedoch erreichten lediglich die Schüler der Versuchsgruppe LDH im Wissenspretest signifikant mehr Punkte als die entsprechende Kontrollgruppe ($t(195)=2,46$; $p\leq 0,05$; $\omega^2=0,025$). Wie die geringe Effektstärke andeutet, ist der Unterschied allerdings nur wenig bedeutsam. Dennoch wird in Anlehnung an vergleichbare Datenauswertungen in anderen Studien (vgl. Natzke & Petermann 2009) das Pretestwissen der LDH-Gruppe im Rahmen der Längsschnittanalyse als Kovariate berücksichtigt.

Tabelle 113: Mittelwerte (M) und Standardabweichungen (SD) der im Wissenspretest der drei Versuchsgruppen und der entsprechenden Kontrollgruppe erreichten Punktezahl (vgl. Abbildungen 60 bis 62).

	Treatmentgruppe	M	SD	N
Vorwissen	Versuchsgruppe Gen. Fingerprint	31,60	6,47	171
	Kontrollgruppe (Wissenstest Gen. Finger)	33,39	7,97	28
Vorwissen	Versuchsgruppe LDH	25,43	4,03	130
	Kontrollgruppe (Wissenstest LDH)	23,91	4,28	67
Vorwissen	Versuchsgruppe Einzeller	30,65	4,48	75
	Kontrollgruppe (Wissenstest Einzeller)	29,94	3,69	31

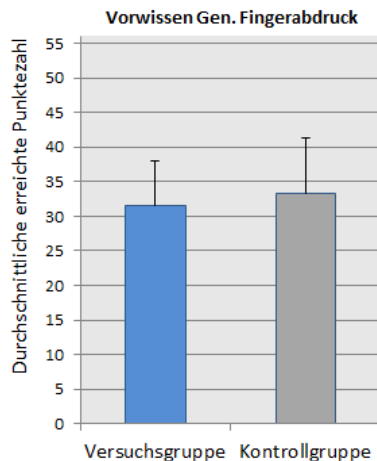


Abbildung 60: Durchschnittliches Vorwissen der Treatmentgruppe Gen. Fingerabdruck sowie der entsprechenden Kontrollgruppe (Skalenmaximum=56 Punkte)

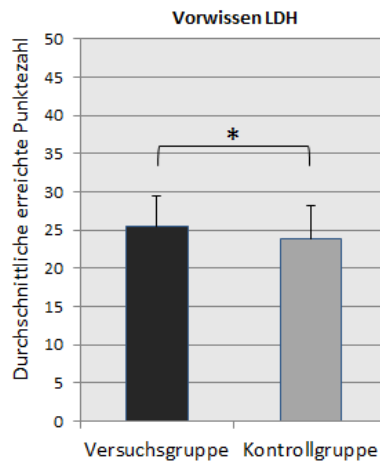


Abbildung 61: Durchschnittliches Vorwissen der Treatmentgruppe LDH sowie der entsprechenden Kontrollgruppe (Skalenmaximum=50 Punkte; * $p \leq 0,05$)

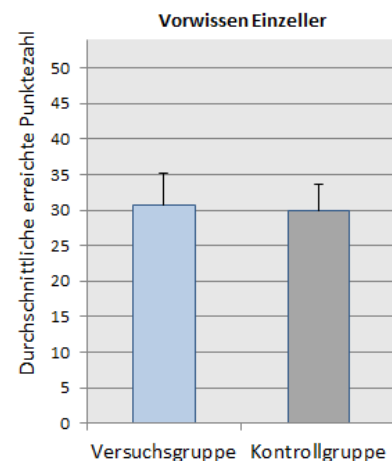


Abbildung 62: Durchschnittliches Vorwissen der Treatmentgruppe Gen. Fingerabdruck sowie der entsprechenden Kontrollgruppe (Skalenmaximum=54 Punkte)

Im vorangegangenen Abschnitt wurden die Zusammenhänge der im Rahmen der Hauptstudie erhobenen dispositionalen Variablen vorgestellt und ein Unterschied im Fachinteresse zwischen den Versuchsgruppen und der Kontrollgruppe aufgedeckt. Weiterhin gelang es, drei in ihrer persönlichen Disposition unterschiedliche Schülertypen zu identifizieren. Diese finden im Rahmen der weiteren Auswertung Berücksichtigung, um typenspezifische Einflüsse des Laborbesuches festzustellen.

3.2.2.2 Laborvariablen

Der folgende Abschnitt stellt Unterschiede in der Wahrnehmung der Laborvariablen zwischen den Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller und der schulischen Kontrollgruppe heraus. Dies geschieht für die Variablen Alltagsbezug, Authentizität, Betreuer, Zusammenarbeit, Druck, Offenheit und Atmosphäre mittels einfaktorieller ANOVAs und sich anschließenden Post-hoc-Tests. Wie bereits erläutert, beziehen sich die Laborvariablen in der Kontrollgruppe nicht auf die Lernumgebung Labor, sondern auf den schulischen Biologieunterricht. Die Kontrollgruppe dient hier somit als externer Standard. Die Variable Anleitung konnte nur in den Laborkursen abgefragt werden, da die Kontrollgruppe „unbehandelt“ blieb und dementsprechend keine Versuchsanleitung eingesetzt wurde.

3.2.2.2.1 Vergleich der Treatmentgruppen

Wie Tabelle 114 und Abbildung 63 zeigen, werden die Variablen Authentizität, Betreuer, Zusammenarbeit, Atmosphäre und Qualität der Versuchsanleitung von den Schülern in den drei Laborkursen sehr hoch eingeschätzt. Insbesondere für die Variablen Authentizität und Atmosphäre

liegen die Mittelwerte der drei Laborkurse deutlich über denen des schulischen Biologieunterrichts. Es deutet sich auch an, dass die Gespräche mit den Betreuern den Schülern im Labor wichtiger sind als die Gespräche mit der Lehrperson im schulischen Biologieunterricht. Druck und Offenheit werden von allen Schülern treatmentunabhängig relativ niedrig eingeschätzt. Ein besonders hoher Alltagsbezug wird im Laborkurs Genetischer Fingerabdruck wahrgenommen. Die durchschnittliche Einschätzung dieses Konstruktes liegt hier deutlich über derjenigen der anderen Versuchsgruppen. Auch die Kontrollgruppe gibt für den schulischen Biologieunterricht einen deutlich geringeren Alltagsbezug an.

Tabelle 114: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Laborvariablen für die drei Laborkurse und die Kontrollgruppe (vgl. Abbildung 63)

	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Alltagsbezug	Versuchsgruppe Gen. Fingerprint	2,74	0,82	218
	Versuchsgruppe Einzeller	2,08	0,93	71
	Versuchsgruppe LDH	2,19	1,03	193
	Kontrollgruppe	2,41	0,69	121
Authentizität	Versuchsgruppe Gen. Fingerprint	3,03	0,68	218
	Versuchsgruppe Einzeller	2,91	0,71	71
	Versuchsgruppe LDH	3,03	0,64	193
	Kontrollgruppe	1,91	0,83	121
Betreuer	Versuchsgruppe Gen. Fingerprint	2,67	0,88	218
	Versuchsgruppe Einzeller	2,89	0,98	71
	Versuchsgruppe LDH	2,78	0,85	193
	Kontrollgruppe	2,40	0,96	121
Zusammenarbeit	Versuchsgruppe Gen. Fingerprint	2,71	0,73	218
	Versuchsgruppe Einzeller	2,86	0,66	71
	Versuchsgruppe LDH	2,71	0,74	193
	Kontrollgruppe	2,73	0,71	121
Druck	Versuchsgruppe Gen. Fingerprint	0,62	0,67	218
	Versuchsgruppe Einzeller	0,85	0,86	71
	Versuchsgruppe LDH	0,65	0,78	193
	Kontrollgruppe	0,98	0,77	121
Offenheit	Versuchsgruppe Gen. Fingerprint	1,06	0,88	218
	Versuchsgruppe Einzeller	1,70	1,06	71
	Versuchsgruppe LDH	1,35	0,82	193
	Kontrollgruppe	1,72	0,87	121
Atmosphäre	Versuchsgruppe Gen. Fingerprint	2,76	0,85	218
	Versuchsgruppe Einzeller	2,76	0,90	71
	Versuchsgruppe LDH	2,77	0,87	193
	Kontrollgruppe	2,10	1,00	121
Anleitung	Versuchsgruppe Gen. Fingerprint	3,24	0,67	224
	Versuchsgruppe Einzeller	3,17	0,76	75
	Versuchsgruppe LDH	3,15	0,64	199

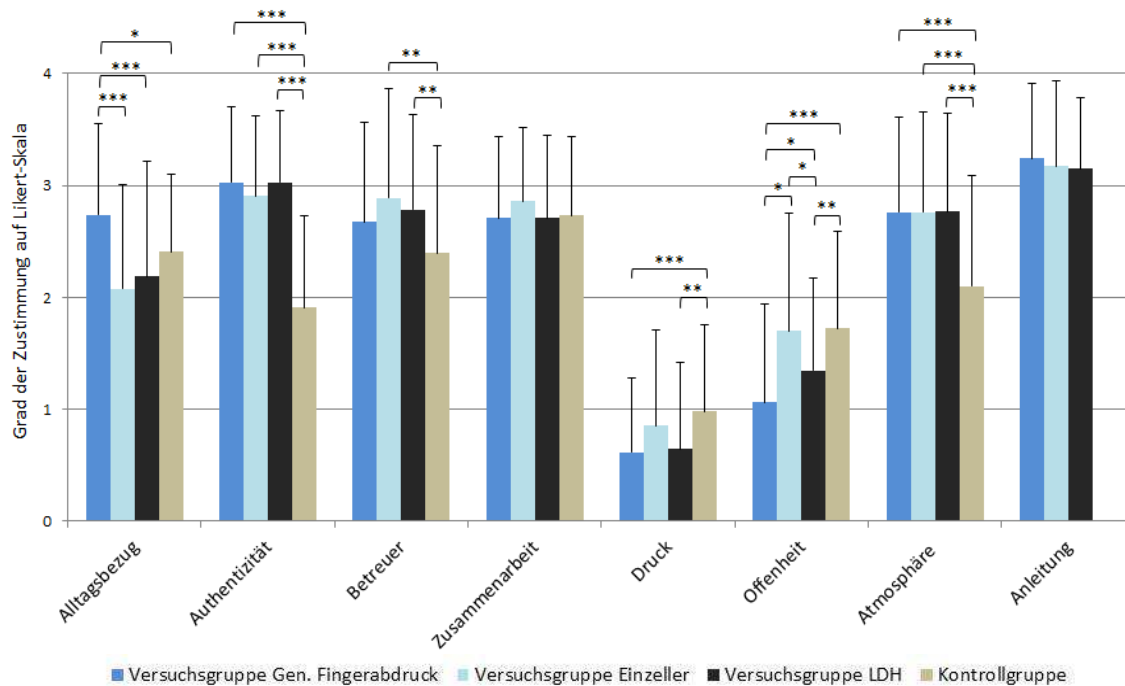


Abbildung 63: Durchschnittliche Zustimmung der Schüler zur Wahrnehmung der Laborvariablen in den drei Laborkursen sowie in der schulischen Kontrollgruppe inklusive Standardabweichungen. Die Variable Anleitung wurde in der Kontrollgruppe nicht erhoben (Skalenmaximum=4; * $p \leq 0,05$; ** $p \leq 0,01$; * $p \leq 0,001$).**

Die durchgeführten ANOVAs bestätigen die augenscheinlichen Beobachtungen statistisch. Für die Laborvariablen Alltagsbezug ($F(3,599)=17,66$; $p \leq 0,001$; $\eta^2=0,081$), Authentizität ($F(3,599)=79,99$; $p \leq 0,001$; $\eta^2=0,286$), Betreuer ($F(3,599)=6,07$; $p \leq 0,001$; $\eta^2=0,029$), Druck ($F(3,599)=7,74$; $p \leq 0,001$; $\eta^2=0,037$), Offenheit ($F(3,599)=19,41$; $p \leq 0,001$; $\eta^2=0,084$) und Atmosphäre ($F(3,599)=17,76$; $p \leq 0,001$; $\eta^2=0,082$) werden höchst signifikante, durch das Treatment bedingte Unterschiede festgestellt. Lediglich für die wahrgenommene Zusammenarbeit sind keine bedeutsamen Unterschiede zu berichten ($F(3,599)=0,68$; $p=0,458$). Auch die wahrgenommene Anleitung unterscheidet sich zwischen den drei Laborkursen nicht ($F(2,494)=0,90$; $p=0,404$). Mit einer aufgeklärten Varianz von 28,6% ($\eta^2=0,286$) ist insbesondere der Effekt der Treatmentgruppen auf die Wahrnehmung der Authentizität als hoch einzustufen. Die anderen Effektstärken liegen im mittleren bis niedrigen Bereich.

Post-hoc-Tests ergeben, dass sich die drei Laborkurse untereinander nicht in der Authentizität unterscheiden. Allerdings liegt die wahrgenommene Authentizität jedes Laborkurses höchst signifikant (Scheffé; $p \leq 0,001$) über der der Kontrollgruppe. Unabhängig von der Art des Laborkurses empfinden die Schüler somit im Schülerlabor eine deutlich höhere Authentizität als im schulischen Biologieunterricht. Weiterhin kann festgestellt werden, dass der empfundene Alltagsbezug der Schüler, die den Laborkurs Genetischer Fingerabdruck besuchten, höchst signifikant (Scheffé; $p \leq 0,001$) über dem der anderen beiden Laborkurse und signifikant (Scheffé; $p \leq 0,05$) über dem empfundenen Alltagsbezug der Kontrollgruppe liegt. Die beiden anderen Laborkurse unterscheiden sich im Alltagsbezug nicht signifikant voneinander und auch nicht von der Kontrollgruppe. Weiterhin

deutet sich an, dass die Schüler im schulischen Biologieunterricht mehr Druck empfinden als in den Schülerlaborkursen. Dies belegen statistisch signifikante Unterschiede zwischen der Kontrollgruppe und den Laborkursen Genetischer Fingerabdruck (Scheffé; $p \leq 0,001$) und LDH (Scheffé; $p \leq 0,01$). Innerhalb der Versuchsgruppen und zwischen dem Kurs Einzeller und der Kontrollgruppe bestehen keine signifikanten Unterschiede im empfundenen Druck. Auch empfinden die Schüler im Biologieunterricht eine höhere Offenheit als in den Kursen Genetischer Fingerabdruck (Scheffé; $p \leq 0,001$) und LDH (Scheffé; $p \leq 0,01$). Der Kurs Einzeller zeichnet sich gegenüber den beiden anderen Laborkursen Genetischer Fingerabdruck (Scheffé; $p \leq 0,001$) und LDH (Scheffé; $p \leq 0,05$) ebenfalls durch eine höhere Offenheit aus. In allen drei Laborkursen empfinden die Schüler die Gespräche mit den betreuenden Personen als gleich wichtig, deutliche Unterschiede sind aber zur Kontrollgruppe festzustellen. Die Schüler der Laborkurse Einzeller (Scheffé; $p \leq 0,01$) und LDH (Scheffé; $p \leq 0,01$) schätzen die Gespräche signifikant, die Teilnehmer des Laborkurses Genetischer Fingerabdruck (Scheffé; $p = 0,62$) tendenziell höher als die Kontrollgruppe ein. Ebenso nehmen die Besucher aller Laborkurse die Atmosphäre im Labor bedeutend positiver wahr (Scheffé; $p \leq 0,001$) als die Atmosphäre im schulischen Biologieraum.

Bisher wurden die drei Laborkurse untereinander und mit der Kontrollgruppe verglichen. Im Anschluss folgt eine differenziertere Betrachtung unter Berücksichtigung des Geschlechts, des Migrationshintergrundes, der beruflichen Orientierung sowie der drei identifizierten Schülertypen. Hierfür werden die drei Laborkurse zusammengefasst.

3.2.2.2 Subgruppenvergleich der Laborwahrnehmung

Um festzustellen, ob die Wahrnehmung der einzelnen Laborkurse beziehungsweise des Biologieunterrichts in der Schule vom Geschlecht, Migrationshintergrund, beruflicher Orientierung oder Schülertypus abhängt, wurden zunächst innerhalb der einzelnen Versuchsgruppen und innerhalb der Kontrollgruppe entsprechende Subgruppenvergleiche vorgenommen. Für die unabhängigen Variablen Geschlecht, Migrationshintergrund und berufliche Orientierung wurden diese anhand von t-Tests durchgeführt. Die Vergleiche des dreigestuften Faktors Schülertypus wurden mittels einfaktoriellen ANOVAs vorgenommen. Vergleiche nach Migrationshintergrund und beruflicher Orientierung sind für die Versuchsgruppe Einzeller nicht möglich, da entsprechende Gruppierungsinformationen hier nicht erhoben wurden.

Statistisch bedeutsame Unterschiede in der Wahrnehmung der Laborvariablen zwischen männlichen und weiblichen Schülern konnten weder in einem der Laborkurse, noch im schulischen Biologieunterricht festgestellt werden. Im Laborkurs Genetischer Fingerabdruck empfanden Schüler

mit Migrationshintergrund eine signifikant höhere Offenheit ($t(54)=2,08$; $p\leq 0,05$; $\omega^2=0,020$) sowie höheren Druck ($t(57)=2,10$; $p\leq 0,05$; $\omega^2=0,020$) als diejenigen ohne Migrationshintergrund. Auch im Laborkurs LDH nimmt diese Subgruppe eine höhere Offenheit wahr ($t(122)=2,11$; $p\leq 0,05$; $\omega^2=0,027$). Außerdem schätzen die Schüler mit Migrationshintergrund die Authentizität des LDH-Kurses etwas höher ein ($t(122)=2,51$; $p\leq 0,05$; $\omega^2=0,041$). Die niedrigen Effektstärken deuten allerdings eine eher geringe Rolle des Migrationshintergrundes in der Wahrnehmung der Laborumgebung an.

Während die Schüler mit naturwissenschaftlicher Orientierung sich in ihrer persönlichen Disposition relativ deutlich von denjenigen ohne naturwissenschaftliche Orientierung abheben, ist in der Laborwahrnehmung in der Versuchsgruppe Genetischer Fingerabdruck kein Unterschied zwischen beiden Subgruppen nachweisbar. Auch im Laborkurs LDH sind die Unterschiede nicht besonders stark ausgeprägt. Dennoch empfinden die naturwissenschaftlich Orientierten hier einen etwas höheren Alltagsbezug ($t(106)=2,61$; $p\leq 0,01$; $\omega^2=0,051$), eine etwas höhere Authentizität ($t(106)=2,01$; $p\leq 0,05$; $\omega^2=0,027$), eine angenehmere Atmosphäre ($t(106)=2,30$; $p\leq 0,05$; $\omega^2=0,038$) und eine höhere Transparenz der Versuchsanleitung ($t(106)=3,36$; $p\leq 0,001$; $\omega^2=0,087$). Die Effektstärken deuten auch in diesem Laborkurs – mit Ausnahme des etwas deutlicheren Effektes in der Variable Versuchsanleitung – eine nur mäßige Bedeutung der beruflichen Orientierung für die Laborwahrnehmung an. Wesentlich ausgeprägter als in den Laborkursen fallen die Unterschiede zwischen naturwissenschaftlich und nicht naturwissenschaftlich orientierten Schülern in der Kontrollgruppe aus. Im schulischen Biologieunterricht empfinden die Schüler mit naturwissenschaftlicher Ausrichtung einen sehr deutlich höheren Alltagsbezug ($t(92)=5,09$; $p\leq 0,001$; $\omega^2=0,209$), eine höhere Authentizität ($t(92)=2,02$; $p\leq 0,05$; $\omega^2=0,032$), eine signifikant höhere Offenheit ($t(92)=2,042$; $p\leq 0,05$; $\omega^2=0,033$) und eine positivere Atmosphäre ($t(92)=3,57$; $p\leq 0,001$; $\omega^2=0,111$) als die nicht naturwissenschaftlich Orientierten. Zudem schätzen diese Schüler die Bedeutung der Gespräche mit der sie betreuenden Biologielehrkraft sehr viel höher ein ($t(92)=3,33$; $p\leq 0,001$; $\omega^2=0,097$) und empfinden ihre Zusammenarbeit mit ihren Mitschülern besser ($t(92)=3,33$; $p\leq 0,001$; $\omega^2=0,097$) als die nicht naturwissenschaftlich Orientierten. Wie die überwiegend im mittleren bis hohen Bereich liegenden Effektstärken andeuten, fällt der Unterschied in der Unterrichtswahrnehmung zwischen naturwissenschaftlich und nicht naturwissenschaftlich orientierten Schülern im schulischen Biologieunterricht deutlicher aus, als dies im Schülerlabor der Fall ist.

Ähnliches ist auch beim Vergleich der drei identifizierten Schülertypen zu beobachten. Auch zwischen den Allroundern, Praktikern und Theoretikern scheinen die Unterschiede im Labor weniger deutlich ausgeprägt zu sein als in der Kontrollgruppe. Inwiefern diese Aussage generalisierbar ist, soll eine

differenzierte Betrachtung der einzelnen Vergleichsgruppen zeigen, die in Abbildung 64 A bis D (Anhang B.2) nachvollzogen werden kann.

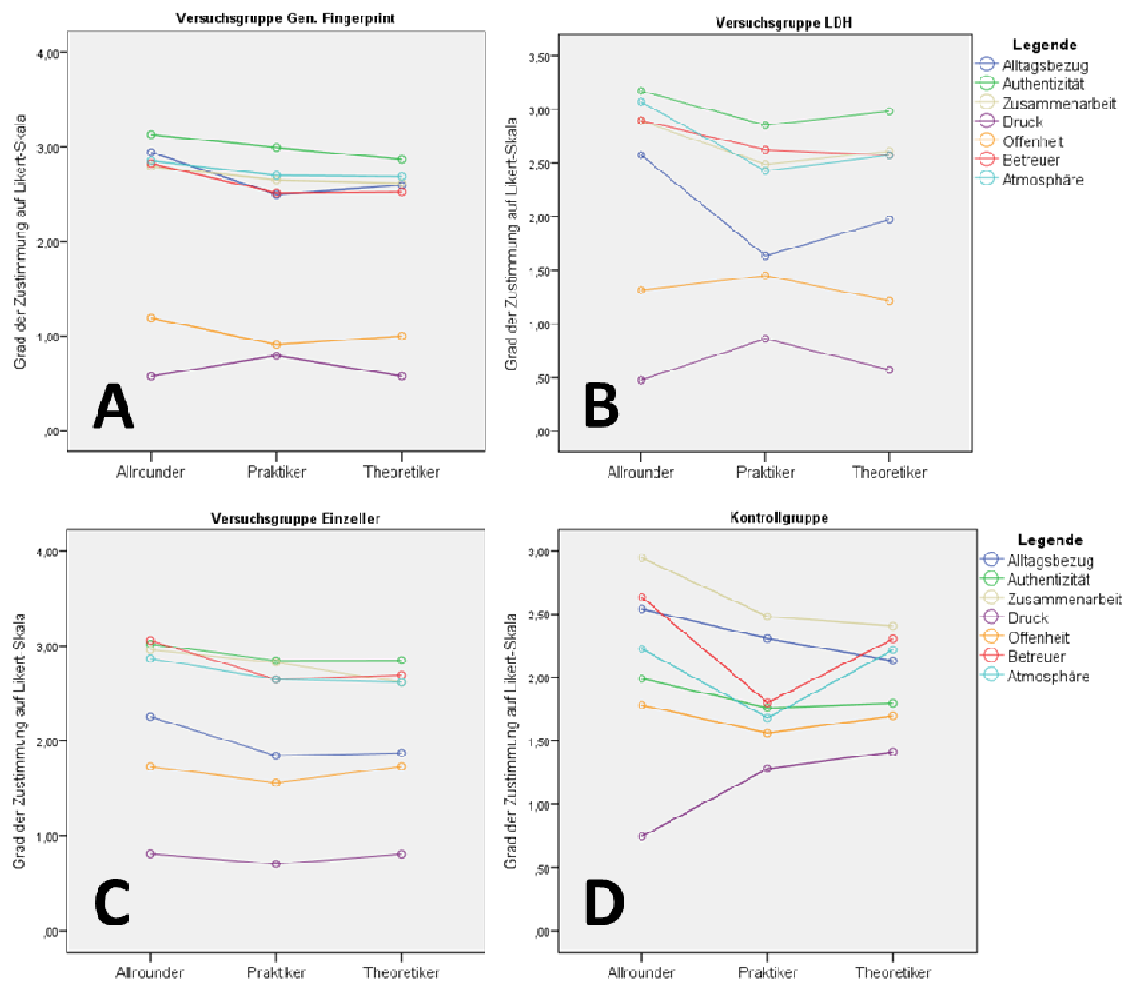


Abbildung 64: Mittelwerte in den Laborvariablen der drei aus der Clusteranalyse hervorgegangenen Gruppen Allrounder, Praktiker und Theoretiker für die Versuchsgruppe Genetischer Fingerabdruck (A), die Versuchsgruppe LDH (B), die Versuchsgruppe Einzeller (C) und die Kontrollgruppe (D) (Skalenmaximum=4).

Der statistische Vergleich der drei Schülertypen mittels einfaktorieller ANOVA zeigt in der Versuchsgruppe Genetischer Fingerabdruck nur relativ geringe Unterschiede zwischen den drei Schülertypen im Alltagsbezug ($F(2,209)=6,35$; $p \leq 0,01$; $\eta^2=0,057$), in der empfundenen Authentizität ($F(2,209)=3,25$; $p \leq 0,05$; $\eta^2=0,030$) sowie in der empfundenen Bedeutung des Betreuers ($F(2,209)=3,26$; $p \leq 0,05$; $\eta^2=0,030$). Mittels Post-hoc-Tests bestätigt sich, dass die Allrounder einen signifikant höheren Alltagsbezug empfinden als die Praktiker (Scheffé; $p \leq 0,01$) und die Theoretiker (Scheffé; $p \leq 0,05$). Zudem empfinden die Allrounder im Laborkurs eine signifikant höhere Authentizität (Scheffé; $p \leq 0,05$) und eine tendenziell höhere Relevanz der Gespräche mit dem Betreuer (Scheffé; $p=0,09$) als die Theoretiker. Die geringen bis – im Falle des Alltagsbezuges – mäßigen Effektstärken lassen die beobachteten Unterschiede als nicht sehr bedeutsam erscheinen. Abbildung 64 A zeigt die Wahrnehmung der Laborvariablen im Laborkurs Genetischer Fingerabdruck für die drei Schülertypen. Ein ähnliches Bild ergibt sich für die Versuchsgruppe Einzeller. In dieser

können keine statistisch bedeutsamen Unterschiede in der Wahrnehmung der Laborvariablen festgestellt werden (vgl. Abbildung 64 C).

Deutlichere Unterschiede sind im Laborkurs LDH zu beobachten. Hier bestätigen einfaktorielle Varianzanalysen signifikante Unterschiede im Alltagsbezug ($F(2,179)=18,56$; $p\leq 0,001$; $\eta^2=0,172$), in der empfundenen Authentizität ($F(2,179)=4,70$; $p\leq 0,01$; $\eta^2=0,050$), in der wahrgenommenen Atmosphäre ($F(2,179)=11,85$; $p\leq 0,001$; $\eta^2=0,117$), im empfundenen Druck ($F(2,179)=5,20$; $p\leq 0,01$; $\eta^2=0,055$) sowie in der wahrgenommenen Qualität der Zusammenarbeit mit den Mitschülern ($F(2,179)=5,88$; $p\leq 0,01$; $\eta^2=0,062$). Im Alltagsbezug sowie in der empfundenen Atmosphäre sind die Unterschiede auf signifikant höhere Einschätzungen der Allrounder gegenüber den Theoretikern und Praktikern zurückzuführen (Scheffé jeweils mindestens $p\leq 0,05$). Zudem finden die Allrounder den Laborkurs LDH bedeutend authentischer und sie empfinden weniger Druck als die Praktiker (Scheffé; $p\leq 0,01$). Weiterhin geben die Allrounder an, mehr mit ihren Mitschülern zusammengearbeitet und kommuniziert zu haben als die Praktiker (Scheffé; $p\leq 0,01$). Während die meisten beobachteten Effekte im Bereich mittlerer Stärke liegen, fällt insbesondere auf, wie unterschiedlich die Schülertypen den Alltagsbezug dieses Kurses wahrnehmen (vgl. Abbildung 64 B).

Ähnlich deutliche Unterschiede zwischen den Schülertypen können bei der Kontrollgruppe (vgl. Abbildung 64 D) beobachtet werden. Hier lassen sich varianzanalytisch bedeutsame Unterschiede im Alltagsbezug ($F(2,116)=3,55$; $p\leq 0,05$; $\eta^2=0,058$), der empfundenen Bedeutung der Gespräche mit dem Betreuer ($F(2,116)=7,80$; $p\leq 0,001$; $\eta^2=0,119$), der Zusammenarbeit mit den Mitschülern ($F(2,116)=7,80$; $p\leq 0,001$; $\eta^2=0,121$) und der Höhe des empfundenen Drucks ($F(2,116)=9,95$; $p\leq 0,001$; $\eta^2=0,146$) feststellen. Mittels Post-hoc-Test nach Scheffé wird ein als signifikant höher empfundener Alltagsbezug der Allrounder gegenüber den Theoretikern belegt (Scheffé; $p\leq 0,05$). Wie Abbildung 64 D erahnen lässt, empfinden die Allrounder zudem signifikant weniger Druck im schulischen Biologieunterricht als die Praktiker und Theoretiker (Scheffé jeweils mindestens $p\leq 0,01$). Weiterhin bekunden die Allrounder beim Experimentieren besser mit ihren Mitschülern zusammenzuarbeiten und zu kommunizieren als die beiden anderen Schülertypen (Scheffé jeweils mindestens $p\leq 0,05$). Zudem schätzen die Praktiker die Kommunikation mit dem Biologielehrer auffällig niedrig ein. Der Unterschied zu den Allroundern ist höchst signifikant (Scheffé, $p\leq 0,001$).

3.2.2.2.3 Korrelationen der Laborvariablen

Tabelle 115 gibt einen gemeinsamen Überblick für alle Versuchsgruppen (N=481) über die Korrelationen der Laborvariablen. Zur besseren Übersicht wurden nicht signifikante Korrelationen entfernt. Korrelationen $r\geq 0,3$ wurden farblich markiert. Höhere Korrelationen der Laborvariablen zu

den Variablen Offenheit, Druck und Atmosphäre bestehen nicht. Nennenswert ist allerdings die positive Korrelation der Variablen Druck und Offenheit ($r(481)=0,281$; $p\leq 0,001$). Schüler, die eine höhere Offenheit empfinden, fühlen sich auch eher unter Druck gesetzt. Deutlich positiv korreliert die Variable Alltagsbezug mit der Authentizität, mit der Bedeutung des Betreuers, der Qualität der Zusammenarbeit und der Atmosphäre. Zu den gleichen Variablen weist auch die Authentizität deutliche Korrelationen auf. Weiterhin sind deutlich positive Korrelationen zwischen der Zusammenarbeit und der Bedeutung des Betreuers sowie der Zusammenarbeit und der empfundenen Atmosphäre festzustellen.

Tabelle 115: Pearson-Korrelationen der Laborvariablen (N=481, gemeinsame Auswertung aller drei Versuchsgruppen; $r\geq 0,3$ ist farblich markiert; nicht signifikante Korrelationen wurden weggelassen)

		Alltagsbezug	Authentizität	Betreuer	Zusammenarbeit	Druck	Offenheit	Atmosphäre	Anleitung
Alltagsbezug	r Signifikanz	1							
Authentizität	r Signifikanz	0,468 $p\leq 0,001$	1						
Betreuer	r Signifikanz	0,365 $p\leq 0,001$	0,397 $p\leq 0,001$	1					
Zusammenarbeit	r Signifikanz	0,339 $p\leq 0,001$	0,370 $p\leq 0,001$	0,334 $p\leq 0,001$	1				
Druck	r Signifikanz			0,113 $p\leq 0,01$		1			
Offenheit	r Signifikanz	0,116 $p\leq 0,05$	0,137 $p\leq 0,01$	0,200 $p\leq 0,001$	0,197 $p\leq 0,001$	0,281 $p\leq 0,001$	1		
Atmosphäre	r Signifikanz	0,373 $p\leq 0,001$	0,313 $p\leq 0,001$	0,296 $p\leq 0,001$	0,304 $p\leq 0,001$	-0,132 $p\leq 0,01$	0,090 $p\leq 0,05$	1	
Anleitung	r Signifikanz	0,208 $p\leq 0,001$	0,132 $p\leq 0,01$		0,243 $p\leq 0,001$	-0,168 $p\leq 0,001$		0,134 $p\leq 0,01$	1

Vorangehend wurde die Wahrnehmung der Lernumgebung Labor durch die Versuchsgruppen dargestellt. Alle bisherigen Analysen beschränkten sich auf singuläre Messzeitpunkte. Im folgenden Abschnitt werden Variablen berücksichtigt, welche über mehrere Messzeitpunkte erhoben wurden.

3.2.2.3 Längsschnittanalysen

Die Variablen Wissen und Fähigkeitsselbstkonzept wurden im Rahmen der Hauptstudie zu allen drei Messzeitpunkten erhoben. Zur Erhebung der experimentbezogenen Selbstwirksamkeitserwartung

wurde im Pre- und Follow-up-Test der FSWEx.V1 eingesetzt. Im Posttest kam der postFSWEx.V1 zum Einsatz, um die unmittelbar in der Experimentiersituation empfundene Selbstwirksamkeit zu erfassen. Unmittelbar nach dem Laborbesuch sowie im Follow-up-Test wurde das aktuelle Interesse erfasst. Die genannten Variablen werden daher im Längsschnitt analysiert.

3.2.2.3.1 Wissen

Aufgrund der unterschiedlichen fachlichen Inhalte der drei evaluierten Laborkurse muss das Fachwissen an dieser Stelle für jeden Kurs einzeln betrachtet werden. Für sämtliche der nachfolgend berichteten Vergleiche wurde getestet, ob bereits zum Zeitpunkt des Pretests Wissensunterschiede bestanden. In diesen Fällen wurde das Vorwissen als Kovariate in der Varianzanalyse berücksichtigt.

3.2.2.3.1.1 Wissen – Genetischer Fingerabdruck

Wie Abbildung 65 zeigt, ist bei den Schülern, welche den Laborkurs besuchten, eine Wissenszunahme vom Pre- zum Posttest sowie ein Rückgang des Wissens vom Post- zum Follow-up-Test festzustellen, während das Wissen der Kontrollgruppe sich über die drei Messzeitpunkte augenscheinlich nicht verändert. Eine einfaktorielle ANOVA mit Messwiederholung bestätigt keinen signifikanten Haupteffekt des Faktors Zeit, also keine zeitliche Veränderung im Wissen der Kontrollgruppe ($F(2,54)=1,547$; $p=0,222$). Ein durch das Ausfüllen des Wissenspretests bedingter Lerneffekt ist daher auszuschließen. Ein Unterschied im Vortestwissen zwischen Versuchs- und Kontrollgruppe ist ebenfalls nicht festzustellen ($t(33)=1,13$; $p=0,267$). Eine zweifaktorielle ANOVA mit Messwiederholung über alle Messzeitpunkte bestätigt eine höchst signifikante Interaktion Zeit x Vergleichsgruppe ($F(2,394)=45,64$; $p\leq 0,001$; $\eta^2=0,188$). Die Versuchsgruppe unterscheidet sich somit in der Wissensentwicklung signifikant von der Kontrollgruppe.

Tabelle 116: Von Versuchs- und Kontrollgruppe im Wissenstest Genetischer Fingerabdruck durchschnittlich erreichte Punktezahl (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*)

Testzeitpunkt	Treatment	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Versuchsgruppe	31,60	6,47	171
	Kontrollgruppe	33,39	7,97	28
Posttest	Versuchsgruppe	41,97	7,15	171
	Kontrollgruppe	31,79	9,29	28
Follow-up-Test	Versuchsgruppe	36,68	7,18	171
	Kontrollgruppe	32,71	8,20	28

Zwischen welchen Messzeitpunkten die Unterschiede verortet sind, zeigen weitere ANOVAs zwischen Pre- und Post-, sowie Pre- und Follow-up-Test. Vom Pre- zum Posttest ist eine höchst signifikante Interaktion der Aspekte Zeit x Gruppe ($F(1,197)=78,20$; $p\leq 0,001$; $\eta^2=0,284$) in Form eines

höchst bedeutsamen Wissenszuwachs der Versuchsgruppe gegenüber der Kontrollgruppe zu beobachten. Dieser ist auch im Vergleich Pre-/Follow-up-Test nachhaltig noch belegbar ($F(1,197)=19,22$; $p \leq 0,001$; $\eta^2=0,089$). Der beobachtete kurz- und langfristige Wissenszuwachs ist somit auf den Laborbesuch zurückzuführen.

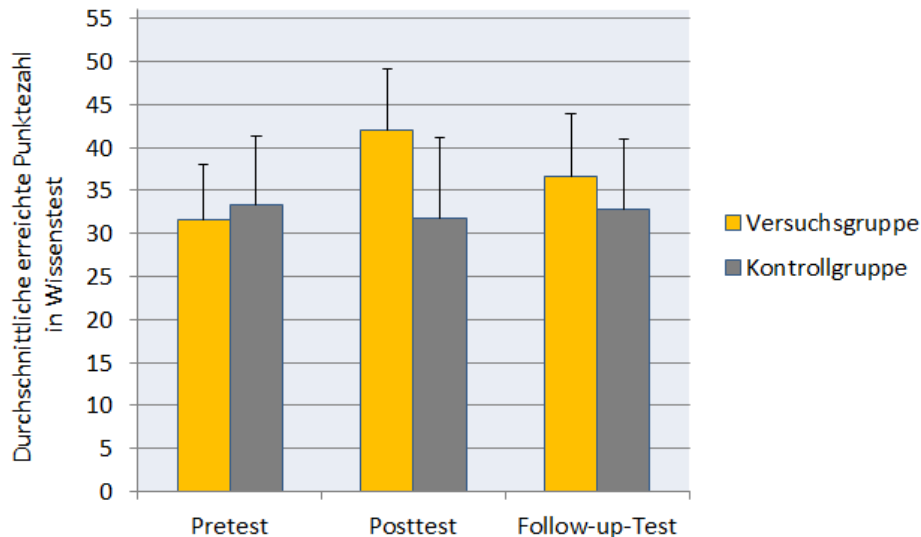


Abbildung 65: Ergebnisse des Wissenstests im Laborkurs Genetischer Fingerabdruck für die Versuchs- und Kontrollgruppe (Skalenmaximum=56 Punkte)

Für weibliche und männliche Schüler, sowie Schüler mit und ohne Migrationshintergrund sind in der Versuchsgruppe weder im Pretest (Geschlecht: $t(169)=0,95$; $p=0,345$); Migrationshintergrund: Geschlecht: $t(169)=0,64$; $p=0,524$), noch in der zeitlichen Veränderung des Wissens (Interaktion Zeit x Geschlecht: $F(2,338)=0,01$; $p=0,992$, Interaktion Zeit x Migrationshintergrund: $F(2,338)=2,03$; $p=0,133$) Unterschiede festzustellen. Unabhängig von Geschlecht und Migrationshintergrund lernen die Schüler somit gleich gut. Schüler mit naturwissenschaftlicher beruflicher Orientierung weisen bereits im Pretest ein signifikant höheres Vorwissen gegenüber den nicht naturwissenschaftlich orientierten auf ($t(153)=4,44$; $p \leq 0,001$; $\omega^2=0,108$), daher wurde eine ANCOVA mit den im Pretest erreichten Wissenspunkten als Kovariate und den Post- und Follow-up-Test-Ergebnissen als abhängige Variablen durchgeführt. Hier ist keine bedeutsame Abhängigkeit der Testergebnisse von der Berufsorientierung festzustellen (Interaktion Zeit x Berufsorientierung: $F(1,152)=5,33$; $p=0,554$). Naturwissenschaftlich und nicht naturwissenschaftlich Orientierte lernen im BeLL Bio-Kurs Genetischer Fingerabdruck also gleich gut. Mittels einfaktorieller univariater ANOVA wurden die Schülertypen der Versuchsgruppe auf Pretestdifferenzen untersucht (vgl. Tabelle 117 & Abbildung 66). Hier konnten statistisch bedeutsame Unterschiede im Vorwissen festgestellt werden ($F(2,165)=5,82$; $p \leq 0,01$; $\eta^2=0,066$), die mittels Post-Hoc-Tests auf ein signifikant höheres Vorwissen der Allrounder gegenüber den Praktikern zurückzuführen sind (Scheffé; $p \leq 0,01$). Das Vorwissen findet daher als Kovariate in den einfaktoriellen ANCOVAs für Post- und Follow-up-Testleistung Anwendung. Ein signifikanter Einfluss des Schülertypus auf den im Posttest erreichten Wissensstand

ist nicht festzustellen (Haupteffekt Schülertypus: $F(2,165)=2,59$; $p=0,078$). Allerdings hat der Schülertypus signifikanten Einfluss auf den langfristigen Wissenszuwachs zum Zeitpunkt des Follow-up-Tests (Haupteffekt Schülertypus: $F(2,164)=4,06$; $p\leq 0,05$; $\eta^2=0,047$). Dieser ist auf einen signifikant höheren, nachhaltigen Wissenszuwachs der Allrounder gegenüber den Praktikern zurückzuführen (Bonferroni, $p\leq 0,05$). Wie die geringe Effektstärke andeutet, ist die Abhängigkeit des Wissenszuwachses vom Schülertypus allerdings als moderat einzustufen.

Tabelle 117: Von den Allroundern, Praktikern und Theoretikern im Wissenstest Genetischer Fingerabdruck durchschnittlich erreichte Punktezahl (M), Standardabweichungen (SD) und Stichprobengrößen (N)

Testzeitpunkt	Schülertypus	M	SD	N
Pretest	Allrounder	32,86	6,33	83
	Praktiker	28,32	4,85	31
	Theoretiker	31,39	7,01	54
Posttest	Allrounder	42,18	6,54	83
	Praktiker	38,65	7,49	31
	Theoretiker	43,26	7,48	54
Follow-up-Test	Allrounder	38,34	6,94	83
	Praktiker	32,29	5,80	31
	Theoretiker	36,24	7,03	54

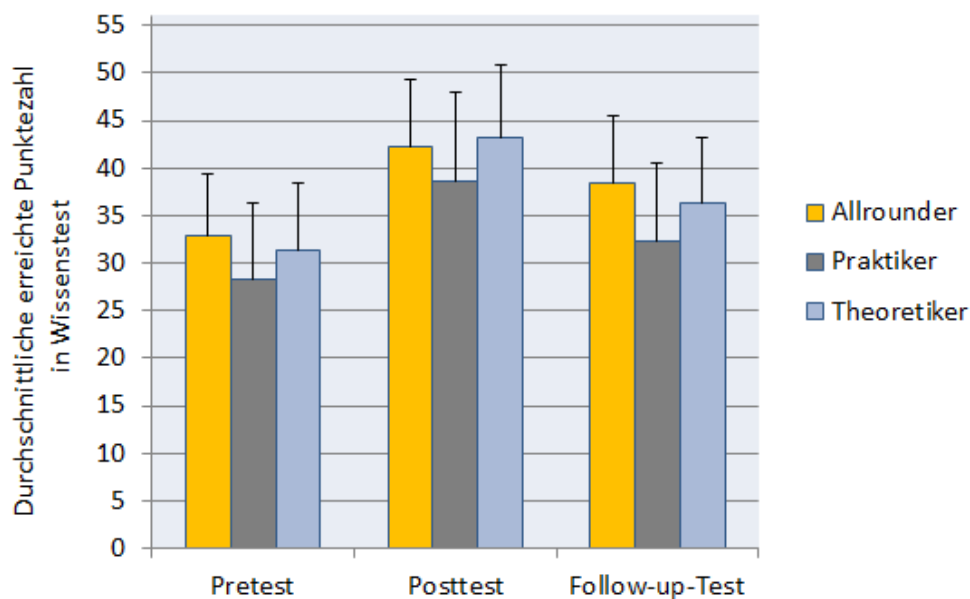


Abbildung 66: Ergebnisse des Wissenstests zum Laborkurs Genetischer Fingerabdruck für die Schülertypen Allrounder, Praktiker und Theoretiker (Skalenmaximum=56 Punkte)

3.2.2.3.1.2 Wissen – LDH

Für die Schüler, welche den Wissenstest LDH bearbeiteten, liegt zum Zeitpunkt des Pretests ein signifikanter Unterschied im Vorwissen von Versuchs- und Kontrollgruppe vor ($t(195)=2,46$, $p\leq 0,05$; $\omega^2=0,025$). Aus diesem Grund wird das Vorwissen als Kovariate im Vergleich des Lernzuwachses der

Treatmentgruppen berücksichtigt. Eine Wissensveränderung der Kontrollgruppe über die drei Messzeitpunkte ist nicht zu beobachten (Haupteffekt Zeit: $F(2,132)=0,51$; $p=0,604$), ein allein durch das Ausfüllen des Pretests bedingter Lerneffekt ist daher auszuschließen. Sowohl vom Pre- zum Posttest (Haupteffekt Treatment: $F(1,194)=120,99$; $p\leq 0,001$, $\eta^2=0,384$), als auch vom Pre- zum Follow-up-Test (Haupteffekt Treatment: $F(1,194)=65,20$; $p\leq 0,001$, $\eta^2=0,252$) lernt die Versuchsgruppe höchst signifikant gegenüber der Kontrollgruppe dazu. Auch im Laborkurs LDH bestätigt sich somit ein nachhaltiger, durch den Laborbesuch bedingter Wissenserwerb (vgl. Tab. 118 & Abb. 67).

Tabelle 118: Von Versuchs- und Kontrollgruppe im Wissenstest LDH durchschnittlich erreichte Punktezahlen (M), Standardabweichungen (SD) und Stichprobengrößen (N)

	Treatment	M	SD	N
Pretest	Versuchsgruppe	25,43	4,03	130
	Kontrollgruppe	23,91	4,28	67
Posttest	Versuchsgruppe	34,12	6,57	130
	Kontrollgruppe	24,33	3,91	67
Follow-up-Test	Versuchsgruppe	30,88	5,58	130
	Kontrollgruppe	24,36	3,88	67

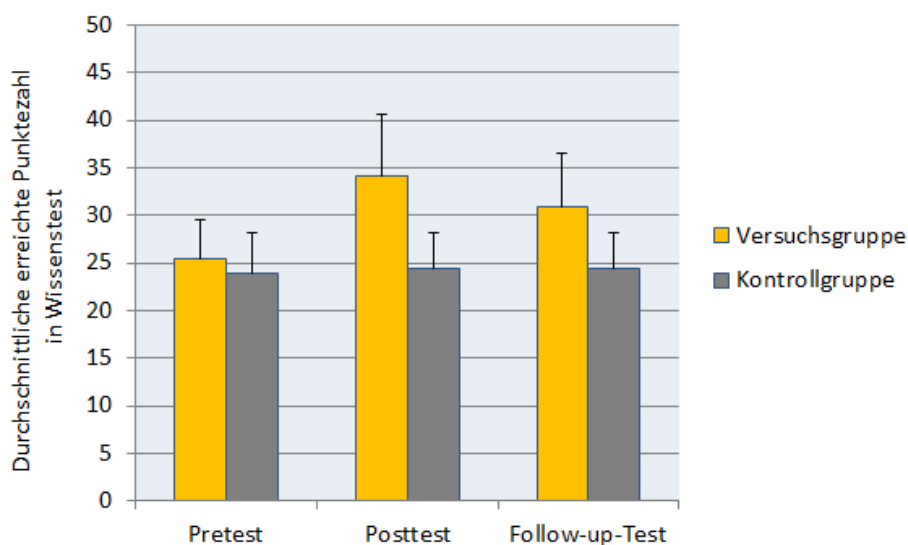


Abbildung 67: Ergebnisse des Wissenstests im Laborkurs LDH für die Versuchs- und Kontrollgruppe (Skalenmaximum=50 Punkte)

Innerhalb der Versuchsgruppe sind zwischen männlichen und weiblichen Schülern, zwischen Schülern mit und ohne Migrationshintergrund sowie naturwissenschaftlich und nicht naturwissenschaftlich Orientierten, weder im Pretest (Geschlecht: $t(128)=1,86$, $p=0,067$; Migrationshintergrund: $t(128)=0,05$, $p=0,964$; Berufliche Orientierung: $t(112)=0,87$, $p=0,387$), noch in der Wissensveränderung statistisch bedeutsame Unterschiede über die drei Messzeitpunkte festzustellen (Interaktion Zeit x Geschlecht: $F(2,256)=1,42$; $p=0,245$; Interaktion Zeit x Migrationshintergrund: $F(2,256)=0,37$; $p=0,694$; Interaktion Zeit x Berufliche Orientierung:

$F(2,224)=1,92$; $p=0,150$). Im Gegensatz zum Kurs Genetischer Fingerabdruck bestehen im LDH-Kurs auch zwischen den Allroundern, Praktikern und Theoretikern keine statistisch bedeutsamen Unterschiede im Vortestwissen (Haupteffekt Schülertyp: $F(2,124)=0,24$; $p=0,791$). Es ist zwar zu beobachten, dass Allrounder und Theoretiker kurz- und langfristig etwas mehr dazu lernen als die Praktiker, dieser Unterschied ist aber bei Einbeziehung aller drei Messzeitpunkte in die ANOVA mit Messwiederholung nicht signifikant (Interaktion Zeit x Schülertyp: $F(4,248)=1,55$; $p=0,189$).

3.2.2.3.1.3 Wissen – Einzeller

Für den Laborkurs Einzeller liegen keine Follow-up-Test-Daten vor, so dass hier nur das Wissen in Pre- und Posttest berichtet werden kann. Weiterhin ist keine Differenzierung nach Migrationshintergrund und beruflicher Orientierung möglich, da diese Daten erst im Follow-up-Test erfragt wurden. Versuchs- und Kontrollgruppe unterscheiden sich zum Zeitpunkt des Pretests nur unwesentlich im Vorwissen ($t(104)=0,79$, $p=0,433$). Eine Veränderung des Wissensstandes vom Pre- zum Posttest ist in der Kontrollgruppe nicht zu beobachten, so dass ein Pretesteffekt auszuschließen ist (Haupteffekt Zeit: $F(1,30)=0,10$; $p=0,758$). Vom Pre- zum Posttest kann mittels zweifaktorieller ANOVA mit Messwiederholung ein gegenüber der Kontrollgruppe höchst signifikanter Lernzuwachs der Versuchsgruppe festgestellt werden (Interaktion Zeit x Treatment: $F(1,104)=62,04$; $p\leq 0,001$; $\eta^2=0,374$). Auch in diesem Laborkurs lernen die Schüler somit – bedingt durch den Laborbesuch – höchst signifikant dazu (vgl. Tab. 119 & Abb. 68).

Tabelle 119: Von Versuchs- und Kontrollgruppe im Wissenstest Einzeller durchschnittlich erreichte Punktezahlen (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*)

	Treatment	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Versuchsgruppe	30,65	4,48	75
	Kontrollgruppe	29,94	3,69	31
Posttest	Versuchsgruppe	41,01	5,62	75
	Kontrollgruppe	29,68	5,24	31

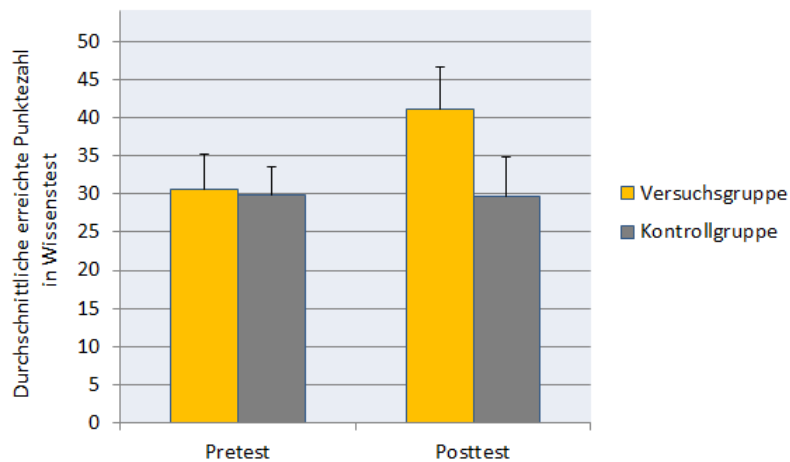


Abbildung 68: Ergebnisse des Wissenstests im Laborkurs Einzeller für die Versuchs- und Kontrollgruppe (Skalenmaximum=54 Punkte)

Beim Vergleich der Geschlechter innerhalb der Versuchsgruppe sind ein geringfügiger Wissensvorsprung der weiblichen Schüler im Pretest sowie ein augenscheinlich etwas höherer Lernzuwachs der männlichen Schüler beobachtbar. Der Unterschied im Vorwissen ist signifikant ($t(73)=2,07$, $p \leq 0,05$; $\omega^2=0,042$), so dass das Wissen als Kovariate in die Varianzanalyse einfließt. Hier wird lediglich ein tendenziell höherer Wissenszuwachs der männlichen Schüler bestätigt (Haupteffekt Geschlecht: $F(1,72)=2,91$; $p=0,092$). Im Vorwissen (Haupteffekt Schülertyp: $F(2,67)=0,14$; $p=0,871$) sowie im Lernzuwachs der Praktiker, Theoretiker und Allrounder (Interaktion Zeit x Schülertyp: $F(2,67)=0,61$; $p=0,549$) ist kein Unterschied auszumachen.

3.2.2.3.2 Fähigkeitsselbstskonzept

In den Versuchsgruppen Genetischer Fingerabdruck und LDH sowie in der Kontrollgruppe wurde das Fähigkeitsselbstskonzept zu allen drei Messzeitpunkten erhoben. Da für den Einzeller-Kurs kein Follow-up-Test erfolgte, werden hier nur Beobachtungen im Pre- und Posttest berichtet. Tabelle 120 & Abbildung 69 stellen die Höhe des biologischen Fähigkeitsselbstkonzeptes der Schüler zu jedem der drei Messzeitpunkte dar.

Für die Kontrollgruppe kann mittels einfaktorieller Varianzanalyse mit Messwiederholung über die drei Messzeitpunkte keine signifikante Veränderung des Fähigkeitsselbstkonzeptes festgestellt werden (Haupteffekt Zeit: $F(2,202)=1,76$; $p=0,174$).

Tabelle 120: Mittelwerte (*M*) und Standardabweichungen (*SD*) des Fähigkeitsselbstkonzeptes der drei Versuchsgruppen und der Kontrollgruppe im Pre-, Post- und Follow-up-Test inklusive Angabe der Stichprobengrößen (*N*)

Testzeitpunkt	Versuchsgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Genetischer Fingerabdruck	2,87	0,91	169
	LDH	2,65	1,09	130
	Einzeller	2,73	0,99	72
	Kontrollgruppe	2,89	0,89	102
Posttest	Genetischer Fingerabdruck	3,09	0,84	169
	LDH	2,86	1,09	130
	Einzeller	2,89	1,01	72
	Kontrollgruppe	2,91	0,93	102
Follow-up-Test	Genetischer Fingerabdruck	2,91	0,91	169
	LDH	2,66	1,10	130
	Kontrollgruppe	2,78	0,99	102

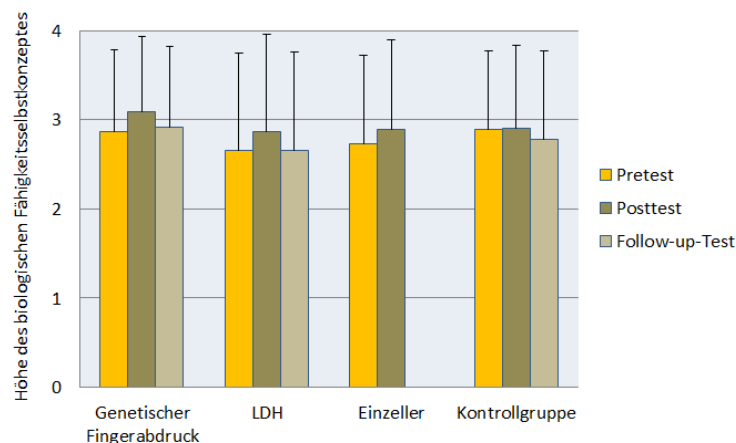


Abbildung 69: Entwicklung des Fähigkeitsselbstkonzeptes im Pre-, Post- und Follow-up-Test für die drei Versuchsgruppen und die Kontrollgruppe (Skalenmaximum=4)

Beobachtete Veränderungen bei den Versuchsgruppen sind somit auf das Labortreatment zurückzuführen. Das Fähigkeitsselbstkonzept der Schüler in der Versuchsgruppe Genetischer Fingerabdruck unterscheidet sich zum Zeitpunkt des Pretests nicht relevant von dem der Kontrollgruppe ($t(269)=0,14$; $p=0,889$). Eine zweifaktorielle Varianzanalyse mit Messwiederholung belegt aber einen signifikanten Unterschied in der zeitlichen Änderung des Fähigkeitsselbstkonzeptes im Vergleich zur Kontrollgruppe (Interaktion Zeit x Treatment: $F(2,538)=3,11$; $p\leq 0,05$; $\eta^2=0,011$). Im Rahmen einer einfaktoriellen ANOVA mit Messwiederholung über Pre- und Posttest kann dieser Unterschied auf eine signifikante, kurzfristige Steigerung des Fähigkeitsselbstkonzeptes in der Versuchsgruppe Genetischer Fingerabdruck gegenüber der Kontrollgruppe zurückgeführt werden (Interaktion Zeit x Treatment: $F(1,269)=5,80$; $p\leq 0,05$; $\eta^2=0,021$). Langfristig geht dieses Selbstkonzept aber wieder fast auf Pretestniveau zurück, so dass vom Pre- zum Follow-up-Test keine treatmentbedingte Steigerung mehr festzustellen ist (Interaktion Zeit x Treatment: $F(1,269)=3,024$; $p=0,083$). Ähnliches ist in der Versuchsgruppe LDH zu beobachten. Auch hier ist zum Zeitpunkt des

Pretests kein bedeutsamer Unterschied im Fähigkeitsselbstkonzept zur Kontrollgruppe nachweisbar ($t(230)=1,75$; $p=0,081$). Augenscheinlich nimmt auch hier das Fähigkeitsselbstkonzept kurzfristig laborkursbedingt zu. Allerdings stellt sich dieser Unterschied mittels ANOVA mit Messwiederholung über die drei Messzeitpunkte als nicht signifikant heraus (Interaktion Zeit x Treatment: $F(1,460)=1,82$; $p=0,163$). Auch im Kurs Einzeller unterscheidet sich das Fähigkeitsselbstkonzept zum Zeitpunkt Pretest nicht bedeutsam von der Kontrollgruppe ($t(172)=1,08$; $p=0,284$). Wie im Laborkurs LDH deutet sich auch hier augenscheinlich eine Erhöhung des Fähigkeitsselbstkonzeptes unmittelbar nach dem Laborkurs an, diese ist allerdings nicht statistisch relevant (Interaktion Zeit x Treatment: $F(1,172)=1,08$; $p=0,299$).

In einer differenzierteren Analyse wurde für jede der Versuchsgruppen untersucht, ob das Geschlecht, der Migrationshintergrund, die berufliche Orientierung oder der Schülertypus Einfluss auf die zeitliche Entwicklung des Fähigkeitsselbstkonzeptes nimmt. Für den Laborkurs Einzeller liegen keine Informationen über Migrationshintergrund und berufliche Orientierung der Schüler vor, so dass hier lediglich das Geschlecht und der Schülertypus berücksichtigt werden. In keinem der Kurse ließ sich zum Zeitpunkt des Pretests ein signifikanter Unterschied zwischen männlichen und weiblichen Schülern im Fähigkeitsselbstkonzept feststellen (Genetischer Fingerabdruck: $t(167)=1,70$; $p=0,090$; LDH: $t(128)=0,58$; $p=0,563$; Einzeller: $t(128)=1,82$; $p=0,073$). Im Laborkurs Genetischer Fingerabdruck schätzten die männlichen Schüler, im Laborkurs Einzeller die weiblichen Schüler ihr biologisches Fähigkeitsselbstkonzept tendenziell etwas höher ein. Ebenso wirkte sich der Laborbesuch in keinem der Kurse nennenswert unterschiedlich auf das Fähigkeitsselbstkonzept der beiden Geschlechter aus, wie eine ANOVA mit Messwiederholung über die Messzeitpunkte zeigt (Interaktion Zeit x Geschlecht Genetischer Fingerabdruck: $F(2,334)=0,87$; $p=0,421$; LDH: $F(2,256)=0,12$; $p=0,890$; Einzeller: $F(1,70)=2,29$; $p=0,135$).

Weder in der Versuchsgruppe Genetischer Fingerabdruck ($t(167)=0,62$; $p=0,543$), noch in der Versuchsgruppe LDH ($t(77)=1,26$; $p=0,213$) unterscheidet sich zum Zeitpunkt des Pretests das Fähigkeitsselbstkonzept von Schülern mit und ohne Migrationshintergrund. Augenscheinlich erfahren sowohl im Laborkurs Genetischer Fingerabdruck, als auch im LDH-Kurs Schüler ohne Migrationshintergrund eine geringfügig höhere Steigerung des Fähigkeitsselbstkonzeptes als Schüler mit Migrationshintergrund. Die beobachteten Unterschiede sind über die drei Messzeitpunkte weder in der Versuchsgruppe Genetischer Fingerabdruck (Interaktion Zeit x Migrationshintergrund: $F(2,334)=0,61$; $p=0,543$), noch im LDH-Kurs ($F(2,256)=2,28$; $p=0,105$) signifikant.

Deutlichere Differenzen sind dagegen zwischen den Schülern mit und ohne naturwissenschaftlicher beruflicher Orientierung festzustellen. Bereits im Pretest weisen Schüler mit naturwissenschaftlichem Berufsziel sowohl im Kurs Genetischer Fingerabdruck ($t(151)=3,84$; $p\leq 0,001$; $\omega^2=0,083$), als auch im LDH-Kurs ($t(106)=6,20$; $p\leq 0,001$; $\omega^2=0,247$) ein signifikant höheres fachliches

Selbstkonzept auf. Das Fähigkeitsselbstkonzept im Pretest fließt daher als Kovariate in die folgenden einfaktoriellen ANCOVAs ein. In keinem der untersuchten Laborkurse ist ein signifikanter Haupteffekt der beruflichen Orientierung auf die Höhe des Selbstkonzeptes im Posttest (Haupteffekt berufliche Orientierung Genetischer Fingerabdruck $F(1,150)=1,95$; $p=0,164$; LDH: $F(1,111)=0,47$; $p=0,493$) festzustellen. Vom Pre- zum Follow-up-Test zeigen die naturwissenschaftlich orientierten Schüler der Versuchsgruppe Genetischer Fingerabdruck ein noch im Follow-up-Test anhaltend hohes Selbstkonzept, während das der nicht naturwissenschaftlich orientierten Schüler wieder auf Pretestniveau sinkt. Die ANCOVA mit der beruflichen Orientierung als Faktor, dem Selbstkonzept im Pretest als Kovariate, dem Follow-up-Test als abhängige Variable bestätigt einen signifikanten Haupteffekt der beruflichen Orientierung ($F(1,150)=15,54$; $p\leq 0,001$; $\eta^2=0,09$). Im Kurs LDH ist in keiner der Vergleichsgruppen eine nachhaltige Steigerung des Fähigkeitsselbstkonzeptes nachzuweisen (Haupteffekt berufliche Orientierung: $F(1,111)=2,07$; $p=0,153$). Die folgende Analyse der Schülertypen ist von besonderem Interesse. In allen Laborkursen unterscheiden sich zum Zeitpunkt des Pretests die Fähigkeitsselbstkonzepte der Schülertypen nämlich deutlich (Genetischer Fingerabdruck: $F(2,163)=44,05$; $p\leq 0,001$; $\eta^2=0,351$; LDH: $F(2,124)=73,67$; $p\leq 0,001$; $\eta^2=0,543$; Einzeller: $F(2,67)=10,23$; $p\leq 0,001$; $\eta^2=0,234$). Während die Theoretiker und Allrounder im Pretest auf ähnlich hohem Niveau liegen, ist das biologische Fähigkeitsselbstkonzept der Praktiker deutlich geringer als das der beiden anderen Gruppen ausgeprägt, was auch Post-hoc-Tests bestätigen (Scheffé für den Vergleich Praktiker und Allrounder sowie Praktiker und Theoretiker jeweils mindestens: $p\leq 0,05$). Die Pretest-Werte finden daher als Kovariate in den folgenden Kovarianzanalysen Berücksichtigung. Augenscheinlich deutet sich in allen drei Laborkursen eine etwas höhere kurz- und langfristige Steigerung des Fähigkeitsselbstkonzeptes vom Pre- zum Posttest in der Gruppe der Praktiker an. Dies ist in Tabelle 121 und Abbildung 70 für die Versuchsgruppe Genetischer Fingerabdruck und in Tabelle 122 und Abbildung 71 für die Versuchsgruppe LDH dargestellt.

Tabelle 121: Mittelwerte (M) und Standardabweichungen (SD) des Fähigkeitsselbstkonzeptes der drei Schülertypen über die drei Messzeitpunkte Pre-, Post- und Follow-up-Test im Kurs Genetischer Fingerabdruck inklusive Angabe der Stichprobengrößen (N)

Testzeitpunkt	Versuchsgruppe	M	SD	N
Pretest	Allrounder	3,24	0,67	83
	Praktiker	1,82	0,91	31
	Theoretiker	2,94	0,68	52
Posttest	Allrounder	3,43	0,74	83
	Praktiker	2,33	0,90	31
	Theoretiker	3,01	0,64	52
Follow-up-Test	Allrounder	3,26	0,75	83
	Praktiker	2,10	0,94	31
	Theoretiker	2,85	0,80	52

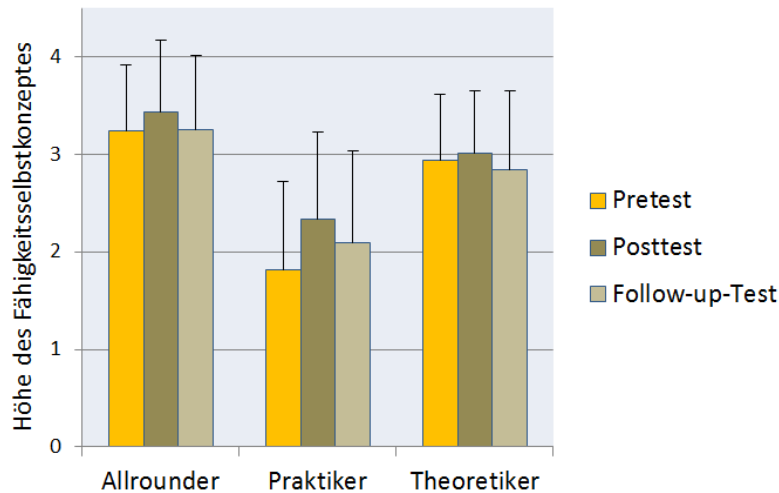


Abbildung 70: Entwicklung des Fähigkeitsselbstkonzeptes der drei Schülertypen Allrounder, Praktiker und Theoretiker über die drei Messzeitpunkte in der Versuchsgruppe Genetischer Fingerabdruck (Skalenmaximum=4)

Tabelle 122: Mittelwerte (*M*) und Standardabweichungen (*SD*) des Fähigkeitsselbstkonzeptes der drei Schülertypen über die drei Messzeitpunkte Pre-, Post- und Follow-up-Test im Kurs LDH inklusive Angabe der Stichprobengrößen (*N*)

Testzeitpunkt	Versuchsgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Allrounder	3,41	0,50	54
	Praktiker	1,62	0,98	47
	Theoretiker	2,99	0,70	26
Posttest	Allrounder	3,49	0,57	54
	Praktiker	1,94	1,05	47
	Theoretiker	3,22	0,86	26
Follow-up-Test	Allrounder	3,37	0,59	54
	Praktiker	1,76	0,95	47
	Theoretiker	2,76	1,03	26

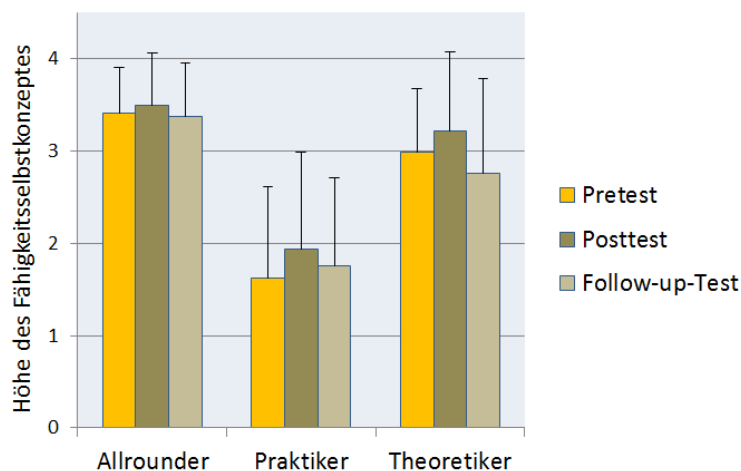


Abbildung 71 Entwicklung des Fähigkeitsselbstkonzeptes der drei Schülertypen Allrounder, Praktiker und Theoretiker über die drei Messzeitpunkte in der Versuchsgruppe LDH (Skalenmaximum=4)

Mittels ANCOVA kann unter Einsatz des Posttests als abhängige Variable und des Pretests als Kovariate in den Kursen Genetischer Fingerabdruck (Haupteffekt Schülertyp: $F(2,1624)=2,19$; $p=0,115$) und Einzeller (Haupteffekt Schülertyp: $F(2,66)=0,00$; $p=0,996$) allerdings keine und im Kurs LDH nur eine tendenziell stärkere kurzfristige Erhöhung des Fähigkeitsselbstkonzeptes der Praktiker gegenüber den Allroundern belegt werden (Haupteffekt Schülertyp: $F(2,123)=2,81$; $p=0,064$; Post-hoc-Test nach Bonferroni $p=0,091$). In den drei Laborkursen unterscheiden sich die Schülertypen zwar untereinander nicht statistisch belegbar in der kurzfristigen Veränderung ihres Fähigkeitsselbstkonzeptes, dafür aber im Falle des Kurses LDH langfristig. Mittels ANCOVA wird in dieser Versuchsgruppe ein signifikanter Haupteffekt des Schülertyps auf die Follow-up-Test-Ergebnisse nachgewiesen (Haupteffekt Schülertyp Genetischer Fingerabdruck: $F(2,162)=1,48$; $p=0,233$; LDH: $F(2,123)=4,33$; $p\leq 0,05$; $\eta^2=0,066$). Dieser ist mittels Post-hoc-Test nach Bonferroni auf eine nachhaltigere Steigerung des Fähigkeitsselbstkonzeptes der Praktiker gegenüber den Allroundern zurückzuführen (Bonferroni; $p\leq 0,05$).

3.2.2.3.3 Aktuelles Interesse

Das aktuelle unmittelbar auf den Laborkurs bezogene Interesse wurde in Post- und Follow-up-Test erhoben. Da für den Kurs Einzeller kein Follow-up-Test durchgeführt wurde, kann an dieser Stelle nur der Längsschnittvergleich für die Kurse Genetischer Fingerabdruck und LDH vorgenommen werden. Da die Kontrollgruppe „unbehandelt“ blieb, entspricht hier das aktuelle Interesse dem generellen, also dem emotionalen, wertbezogenen und epistemischen Interesse am Biologieunterricht. Um einen Vergleichsmaßstab für die Höhe des Interesses in den Laborkursen zeigen zu können, wird es hier als Baseline berücksichtigt. Weil das dispositionale Fachinteresse der Kontrollgruppe deutlich über dem der Versuchsgruppen liegt und da theoriebegründet ein deutlicher Einfluss des persönlichen Fachinteresses auf das aktuelle Interesse anzunehmen ist, wird es in den folgenden Auswertungen als Kovariate berücksichtigt. Tabelle 123 und Abbildung 74 geben für beide Laborkurse und die Kontrollgruppe einen Überblick über die drei erfassten Komponenten des aktuellen Interesses zu den zwei Messzeitpunkten.

Tabelle 123: Mittelwerte (*M*) und Standardabweichungen (*SD*) der drei Komponenten des aktuellen Interesses in den Versuchsgruppen Genetischer Fingerabdruck und LDH sowie der Kontrollgruppe für Post- und Follow-up-Test inklusive Angabe der Stichprobengrößen (*N*)

Skala/Zeitpunkt	Versuchsgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Emotionale Komponente Posttest	Genetischer Fingerabdruck	2,88	0,87	171
	LDH	2,95	0,92	130
	Kontrollgruppe	2,88	0,88	100
Emotionale Komponente Follow-up	Genetischer Fingerabdruck	2,77	0,99	171
	LDH	2,78	0,93	130
	Kontrollgruppe	2,84	0,86	100
Wertebezogene Komponente Posttest	Genetischer Fingerabdruck	2,72	0,79	171
	LDH	2,49	0,88	130
	Kontrollgruppe	2,76	0,75	100
Wertebezogene Komponente Follow-up	Genetischer Fingerabdruck	2,26	0,87	171
	LDH	2,13	0,94	130
	Kontrollgruppe	2,74	0,87	100
Epistemische Komponente Posttest	Genetischer Fingerabdruck	1,81	0,88	171
	LDH	1,67	0,91	130
	Kontrollgruppe	1,99	0,82	100
Epistemische Komponente Follow-up	Genetischer Fingerabdruck	1,36	0,86	171
	LDH	1,29	0,88	130
	Kontrollgruppe	1,93	0,96	100

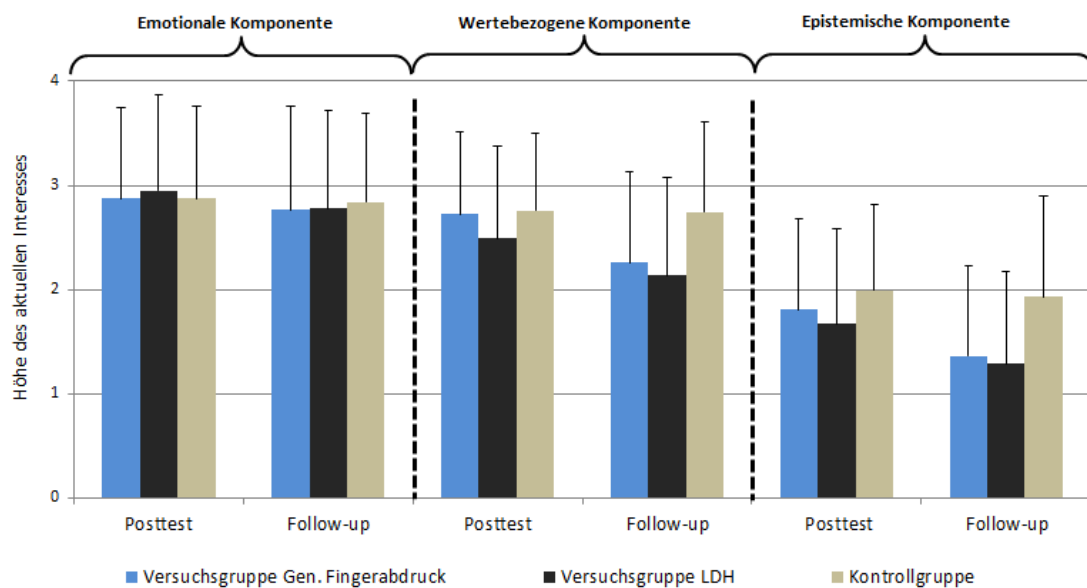


Abbildung 72: Durchschnittliche Höhe der drei Komponenten des aktuellen Interesses im Post- und Follow-up-Test einschließlich der Standardabweichungen (Skalenmaximum=4).

Hier wird deutlich, dass Schüler – unabhängig davon, ob auf schulischen Biologieunterricht oder Schülerlaborkurs bezogen – ihr emotionales und wertbezogenes Interesse deutlich höher einschätzen als ihre epistemische Bereitschaft, sich mit den in Labor oder schulischem Biologieunterricht erlebten Inhalten weiterhin auseinanderzusetzen. Zudem wird ersichtlich, dass das

aktuelle Interesse der Schüler nach der Teilnahme an einem der Laborkurse – also zum Zeitpunkt des Posttests – nicht höher ist als das in der schulischen Kontrollgruppe berichtete Interesse. Außerdem ist für die wertbezogene und die epistemische Interessenkomponente in den Versuchsgruppen ein deutlicher Abfall zum Follow-up-Test hin zu beobachten, während das Interesse der Kontrollgruppe stabil bleibt. Die zeitliche Konstanz der drei Interessenkomponenten wird für diese Gruppe auch statistisch durch einfaktorielle ANOVAs mit Messwiederholung über die drei Messzeitpunkte bestätigt (Haupteffekt Zeit Emotionale Komponente: $F(1,99)=0,38$; $p=0,540$; Epistemische Komponente: $F(1,99)=0,58$; $p=0,445$; Wertebezogene Komponente: $F(1,99)=0,07$; $p=0,797$).

Während das wertbezogene (Haupteffekt Treatment: $F(2,397)=1,61$; $p=0,202$) und das epistemische Interesse (Haupteffekt Treatment: $F(2,397)=0,85$; $p=0,429$) der Versuchsgruppen und der Kontrollgruppe sich im Posttest nicht statistisch bedeutsam unterscheidet, ist überraschenderweise im aktuellen emotionalen Interesse ein eben noch signifikanter Haupteffekt des Treatments festzustellen ($F(2,397)=3,01$; $p\leq 0,05$; $\eta^2=0,015$). Verursacht wird dieses Ergebnis durch einen in allen Interessenkomponenten höchst signifikanten Einfluss der Kovariate Fachinteresse (Haupteffekt Kovariate Epistemisches Interesse: $F(1,397)=57,13$; $p\leq 0,001$; $\eta^2=0,126$; Wertebezogenes Interesse: $F(1,397)=29,39$; $p\leq 0,001$; $\eta^2=0,069$; Emotionales Interesse: $F(1,397)=57,54$; $p\leq 0,001$; $\eta^2=0,127$). Der Einfluss des Fachinteresses auf die Höhe des aktuellen Interesses ist damit sehr deutlich gezeigt. Die Posttest-Unterschiede der Treatmentgruppen in der Interessenkomponente emotionales Interesse sind zwar aufgrund der mäßigen Effektstärke als gering einzustufen. Dennoch wird der Posttest im emotionalen Interesse neben dem dispositionalen Fachwissen in der folgenden ANCOVA als zusätzliche Kovariate berücksichtigt. Hier wird festgestellt, dass das emotionale Interesse der Versuchsgruppen sich ebenso wenig ändert wie das der Kontrollgruppe (Haupteffekt Treatment: $F(2,396)=0,36$; $p=0,697$). Für das epistemische Interesse wird im Rahmen einer ANCOVA mit Messwiederholung (Kovariate Fachwissen) ein höchst signifikanter Unterschied der Treatmentgruppen im Rückgang des epistemischen Interesses festgestellt (Interaktion Zeit x Treatment: $F(2,397)=6,09$; $p\leq 0,01$; $\eta^2=0,030$). Mittels Post-hoc-Test nach Bonferroni kann dieser auf einen signifikanten Rückgang des epistemischen Interesses in den beiden Versuchsgruppen gegenüber der Kontrollgruppe zurückgeführt werden (Bonferroni jeweils $p\leq 0,01$). Im Rahmen einer ANCOVA mit Messwiederholung über den Post- und Follow-up-Test unter Einsatz der Kovariate Fachinteresse wird für das wertbezogene Interesse ebenfalls ein Unterschied zwischen den Treatmentgruppen festgestellt (Interaktion Zeit x Treatment: $F(2,397)=9,25$; $p\leq 0,001$; $\eta^2=0,045$). Mittels Post-hoc-Test nach Bonferroni kann diese Beobachtung auf einen sehr signifikanten Unterschied im Interessenrückgang der Versuchsgruppe LDH und der Kontrollgruppe (Bonferroni; $p\leq 0,01$) sowie auf einen tendenziellen Unterschied zwischen der Versuchsgruppe Genetischer Fingerabdruck und der Kontrollgruppe (Bonferroni; $p=0,71$) zurückgeführt werden. In beiden

Versuchsgruppen nehmen also das aktuelle epistemische sowie das wertebezogene Interesse im Vergleich zur Kontrollgruppe nachweisbar ab, während das emotionale Interesse relativ stabil bleibt.

Weitere Analysen wurden innerhalb der beiden Versuchsgruppen durchgeführt, um festzustellen, ob der beobachtete Interessenrückgang in den Laborkursen vom Geschlecht, Migrationshintergrund, der beruflichen Orientierung oder dem Schülertypus abhängt. Weder im Laborkurs Genetischer Fingerabdruck, noch im Kurs LDH lassen sich unmittelbar nach Besuch der Veranstaltung bedeutsame Unterschiede hinsichtlich der Höhe des aktuellen Interesses zwischen männlichen und weiblichen Schülern finden. Auch im Rückgang des wertebezogenen Interesses (Interaktion Zeit x Geschlecht Genetischer Fingerabdruck: $F(1,168)=1,11$; $p=0,294$; LDH: $F(1,127)=0,082$; $p=0,775$) kann mittels ANCOVA (Kovariate Fachinteresse) mit Messwiederholung kein Unterschied zwischen den Geschlechtern festgestellt werden. Im Kurs Genetischer Fingerabdruck ist allerdings bei den weiblichen Kursteilnehmern ein signifikant höherer Rückgang des epistemischen Interesses zu beobachten als bei den männlichen Schülern (Interaktion Zeit x Geschlecht $F(1,168)=5,10$; $p\leq 0,05$; $\eta^2=0,029$). Ein derartiger Geschlechterunterschied ist im Kurs LDH nicht zu beobachten ($F(1,127)=1,24$; $p=0,267$). Ein signifikanter Unterschied zwischen weiblichen und männlichen Schülern in der Änderung des emotionalen Interesses ist in keinem der Kurse gegeben (Haupteffekt Geschlecht in ANCOVA mit Fachwissen und Posttest als Kovariate; Genetischer Fingerabdruck: $F(1,167)=0,21$; $p=0,649$; LDH: $F(1,126)=3,05$; $p=0,083$).

Während Schüler mit und ohne Migrationshintergrund sich im Kurs LDH kaum unterscheiden, fällt im Kurs Genetischer Fingerabdruck das bereits im Posttest tendenziell geringere ($t(169)=1,81$; $p=0,072$) emotionale sowie das augenscheinlich etwas geringere ($t(169)=1,81$; $p=0,265$) wertebezogene Interesse der Kursteilnehmer mit Migrationshintergrund auf. Unterschiede in der Änderung des wertebezogenen Interesses lassen sich zwischen beiden Schülergruppen nicht feststellen. Offenbar besteht aber in beiden Laborkursen ein etwas stärkerer Rückgang des epistemischen Interesses bei Schülern mit Migrationshintergrund. Dieser Unterschied ist statistisch allerdings nicht zu bestätigen (Interaktion Zeit x Migrationshintergrund: Genetischer Fingerabdruck: $F(1,168)=1,42$; $p=0,235$; LDH: $F(1,127)=2,47$; $p=0,118$). Während das emotionale Interesse im Kurs LDH für Schüler mit und ohne Migrationshintergrund gleichermaßen zum Follow-up-Test hin nur sehr leicht sinkt, ist im Laborkurs Genetischer Fingerabdruck ein signifikant höherer Rückgang dieser Interessenkomponente bei Schülern mit Migrationshintergrund zu beobachten (Haupteffekt Migrationshintergrund: $F(1,167)=4,29$; $p\leq 0,05$; $\eta^2=0,025$).

Weiterhin fällt im Kurs LDH unmittelbar nach dem Laborbesuch ein deutlich niedrigeres emotionales ($t(112)=1,86$; $p=0,065$), epistemisches ($t(112)=2,31$; $p\leq 0,05$; $\omega^2=0,037$) und wertbezogenes Interesse ($t(112)=2,25$; $p\leq 0,05$; $\omega^2=0,034$) der nicht naturwissenschaftlich orientierten Schüler auf. Für das epistemische ($t(153)=1,21$; $p=0,227$) und wertbezogene ($t(153)=1,48$; $p=0,141$) Interesse ist dies zwar auch nach dem Kurs Genetischer Fingerabdruck zu beobachten, allerdings nicht auf statistisch bedeutsamen Niveau. Bedeutende Unterschiede bezüglich des Interessenrückgangs zwischen beiden Gruppen sind nicht nachzuweisen.

In Bezug auf die Schülertypen stellt sich der Allrounder in beiden Laborkursen, in allen Interessenkomponenten und zu beiden Messzeitpunkten als derjenige mit dem höchsten aktuellen Interesse heraus. An zweiter Position folgt ihm im Kurs LDH der Theoretiker, während der Praktiker hier in allen Komponenten das geringste aktuelle Interesse bekundet. Etwas anders ist die Situation in der Versuchsgruppe Genetischer Fingerabdruck. Wie im LDH-Kurs geben die Allrounder auch hier im Posttest das höchste Interesse an, gefolgt von den Theoretikern und – fast gleichauf – den Praktikern. Im epistemischen Interesse und dem wertbezogenen Interesse liegen die Theoretiker jedoch hinter den Praktikern. Signifikante Unterschiede zwischen den Schülertypen sind zum Posttestzeitpunkt allerdings nur für das wertbezogene Interesse in der Versuchsgruppe genetischer Fingerabdruck festzustellen ($F(2,164)=4,04$; $p\leq 0,05$; $\eta^2=0,047$). Mittels Post-hoc-Test nach Bonferroni wird lediglich ein tendenziell höheres Wissen der Allrounder gegenüber den Theoretikern festgestellt (Bonferroni; $p=0,051$). Von den Praktikern unterscheiden sie sich dagegen nicht bedeutsam. Bei allen drei Schülertypen verläuft der Rückgang des aktuellen Interesses zum Follow-up-Test hin in beiden Laborkursen ähnlich. Statistisch bedeutsame Unterschiede sind mittels ANCOVA mit Messwiederholung und dem dispositionalen Fachinteresse als Kovariate in keiner Interessenkomponente zu belegen.

3.2.2.3.4 Selbstwirksamkeit des Experimentierens

In der Hauptstudie wurde der FSWEx.V1 in identischer Form in Pre- und Follow-up-Test eingesetzt, um die langfristige Konstanz der experimentbezogenen Selbstwirksamkeitserwartung zu prüfen. Da für den Kurs Einzeller keine Follow-up-Test-Daten vorliegen, beschränkt sich die Ausführung auf die Versuchsgruppen Genetischer Fingerabdruck und LDH sowie auf die Resultate der Kontrollgruppe. Im Posttest, also unmittelbar nach dem Laborbesuch, wurde die abgewandelte Version postFSWEx.V1 eingesetzt, um die im jeweiligen Laborkurs wahrgenommene Selbstwirksamkeit zu messen. Da die absolute Höhe der eingeschätzten Selbstwirksamkeit von der Formulierung der Items abhängen kann, wurden diese Posttest-Ergebnisse lediglich über Korrelationen mit den FSWEx.V1-Daten abgeglichen. Der postFSWEx.V1 wurde auch in der Versuchsgruppe Einzeller eingesetzt.

3.2.2.3.4.1 FSWEx.V1 – Subskala Planung

Zum Zeitpunkt des Pretests schätzen die beiden Versuchsgruppen sowie die Kontrollgruppe ihre Selbstwirksamkeit, Experimente erfolgreich planen zu können, identisch ein ($F(2,392)=0,757$; $p=0,470$). Eine bedeutsame Veränderung der Selbstwirksamkeitserwartung vom Pre- zum Follow-up-Test ist in keiner der Gruppen zu beobachten, wie eine zweifaktorielle ANOVA mit Messwiederholung bestätigt (Haupteffekt Zeit: $F(1,392)=0,54$; $p=462$; vgl. Tab. 124 und Abb. 73).

Tabelle 124: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie die Stichprobengrößen (*N*) der Treatmentgruppen in der FSWEx.V1-Subskala Planung in Pre- und Follow-up-Test

	Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
FSWEx.V1 – Subskala Planung	Pretest	Genetischer Fingerabdruck	1,67	0,65	168
		LDH	1,73	0,67	125
		Kontrollgruppe	1,77	0,67	102
	Follow-up-Test	Genetischer Fingerabdruck	1,70	0,75	168
		LDH	1,70	0,79	125
		Kontrollgruppe	1,70	0,72	102

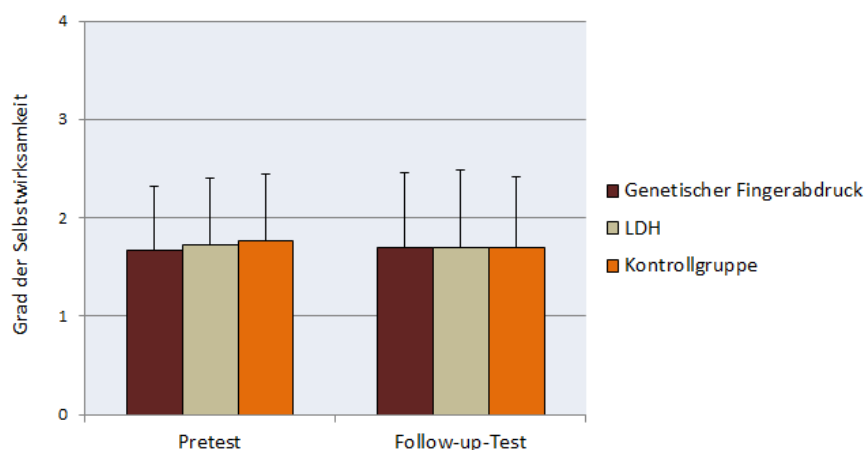


Abbildung 73: Mittelwerte der FSWEx.V1-Subskala Planung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4).

Beim Vergleich der Geschlechter ist festzustellen, dass die männlichen Kursteilnehmer in beiden Versuchsgruppen zum Zeitpunkt des Pretests ihre Planungskompetenz höher einschätzen als die weiblichen Teilnehmer. Während dieser Unterschied für die Versuchsgruppe LDH nur geringfügig ausfällt ($t(123)=0,834$; $p=0,406$), ist er in der Versuchsgruppe Genetischer Fingerabdruck höchst signifikant ($t(166)=3,48$; $p\leq 0,001$; $\omega^2=0,062$). Entsprechend finden die Pretestergebnisse in den Berechnungen für den Kurs Genetischer Fingerabdruck als Kovariate Berücksichtigung. Langfristige Änderungen der Selbstwirksamkeit fallen bei beiden Geschlechtern gering und nicht geschlechtsspezifisch aus. Für Schüler mit und ohne Migrationshintergrund sind in beiden Versuchsgruppen weder im Pretest noch in der zeitlichen Veränderung der Selbstwirksamkeit Unterschiede festzustellen. Schüler mit naturwissenschaftlicher Berufsorientierung schätzen in der Versuchsgruppe Genetischer Fingerabdruck ihre Planungskompetenz zum Pretest-Zeitpunkt sehr

signifikant höher ein als die nicht naturwissenschaftlich orientierten ($t(150)=3,48$; $p\leq 0,01$; $\omega^2=0,056$). Dieser Unterschied ist auch in der Versuchsgruppe LDH zu beobachten, allerdings nicht statistisch bedeutsam ($t(110)=2,35$; $p=0,181$). Zum Follow-up-Test hin ist bei den naturwissenschaftlich Orientierten in der Versuchsgruppe Genetischer Fingerabdruck ein minimaler und im LDH-Kurs ein deutlicherer Anstieg der Selbstwirksamkeitserwartung festzustellen. Die Selbsteinschätzung der nicht naturwissenschaftlich Orientierten verändert sich dagegen im Laborkurs Genetischer Fingerabdruck gar nicht, im LDH-Kurs sinkt sie geringfügig. ANCOVAs mit den Pretestwerten als Kovariate bestätigen in der Versuchsgruppe Genetischer Fingerabdruck einen tendenziellen Anstieg der Selbstwirksamkeitserwartung bei den naturwissenschaftlich Orientierten gegenüber der Vergleichsgruppe (Haupteffekt berufliche Orientierung: $F(1,149)=3,62$; $p=0,059$). Im Laborkurs LDH ist dieser Unterschied sogar sehr signifikant (Haupteffekt berufliche Orientierung: $F(1,106)=9,15$; $p\leq 0,01$; $\eta^2=0,080$). Auch zwischen den Schülertypen sind bereits im Pretest sowohl in der Versuchsgruppe Genetischer Fingerabdruck ($F(2,165)=33,08$; $p\leq 0,001$; $\eta^2=0,286$), als auch im LDH-Kurs ($F(2,124)=19,85$; $p\leq 0,001$; $\eta^2=0,242$) deutliche Selbstwirksamkeitsdifferenzen erkennbar. Diese sind in beiden Versuchsgruppen auf eine deutlich höhere Selbstwirksamkeit der Allrounder gegenüber den Praktikern (Scheffé; $p\leq 0,001$) und den Theoretikern (Scheffé; $p\leq 0,001$) zurückzuführen. In der Versuchsgruppe Genetischer Fingerabdruck liegen die Praktiker etwa gleichauf mit den Theoretikern, während die Theoretiker im LDH-Kurs ihre Planungskompetenz etwas höher einstufen. Signifikante typenabhängige Veränderungen der Selbstwirksamkeit vom Pre- zum Follow-up-Test sind nicht zu beobachten.

3.2.2.3.4.2 postFSWEx.V1 – Subskala Planung

Da der postFSWEx.V1 sich unmittelbar auf die Experimentiertätigkeit in den BeLL Bio-Laborkursen bezieht, kann hier die Kontrollgruppe nicht als Vergleichsgruppe verwendet werden. Hier werden daher ausschließlich die Ergebnisse der Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller berichtet. Ähnlich wie im FSWEx.V1 schätzen Schüler ihre Planungskompetenz, auch bezogen auf konkrete Experimentiersituationen, eher gering ein. Varianzanalytisch ist ein signifikanter Unterschied in der Selbsteinschätzung der Schüler zwischen den Versuchsgruppen festzustellen ($F(2,466)=3,42$; $p\leq 0,05$; $\eta^2=0,014$). Die geringe Effektstärke zeigt, dass dieser Unterschied nicht allzu bedeutsam ist. Auch ein anschließender Post-hoc-Test nach Scheffé kann nur einen tendenziellen Unterschied zwischen den höheren Einschätzungen in der Versuchsgruppe Einzeller und den geringeren Einschätzungen in der Gruppe LDH belegen (Scheffé; $p=0,06$). Tabelle 125 und Abbildung 74 stellen die in den Laborkursen empfundene experimentbezogene Selbstwirksamkeitserwartung in der Subskala Planung dar.

Tabelle 125: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie die Stichprobengrößen (*N*) der Versuchsgruppen in der postFSWEx.V1-Subskala Planung im Posttest.

Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Posttest	Genetischer Fingerabdruck	1,54	0,82	216
	LDH	1,39	0,80	184
	Einzeller	1,66	0,73	69

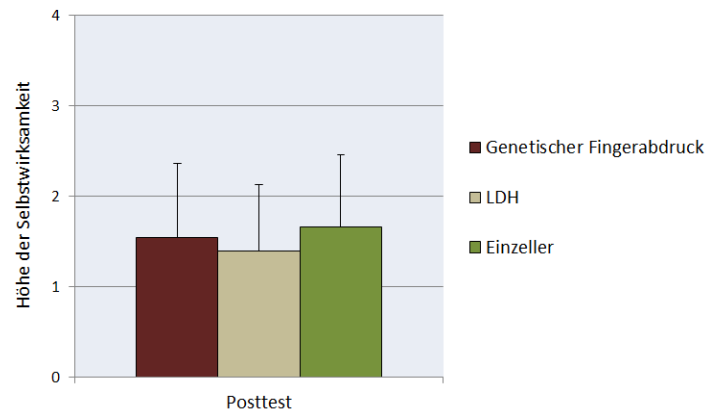


Abbildung 74: Mittelwerte der postFSWEx.V1-Subskala Planung der Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller im Posttest inklusive Standardabweichungen (Skalenmaximum=4)

Subgruppenvergleiche innerhalb der Versuchsgruppen erfolgen im Kurs Einzeller bedingt durch den nicht vorhandenen Follow-up-Test nur für die Einflussfaktoren Geschlecht und Schülertypus. Deutliche Unterschiede zwischen den Geschlechtern sind nur für die Gruppe LDH zu bestätigen. Hier schätzen die männlichen Schüler ihre Kompetenz, den Versuch selbst hätten planen zu können, signifikant höher als die weiblichen Schüler ein ($t(182)=3,48$; $p \leq 0,01$; $\omega^2=0,030$). Der Unterschied ist allerdings – wie die Effektstärke andeutet – relativ gering. Auch die Abhängigkeit der eingeschätzten Planungskompetenz vom Migrationshintergrund ist eher gering. Im Kurs Genetischer Fingerabdruck schätzen Schüler mit Migrationshintergrund ihre Planungskompetenz bezogen auf den Laborkurs signifikant höher ein als ihre Mitschüler ohne entsprechenden Hintergrund ($t(163)=2,19$; $p \leq 0,05$; $\omega^2=0,023$). Auch hier spricht das geringe ω^2 für einen vernachlässigbaren Effekt. Zwischen Schülern mit und ohne naturwissenschaftlicher beruflicher Orientierung ist in der Versuchsgruppe Genetischer Fingerabdruck kein und im LDH-Kurs nur ein geringer Unterschied in der eingeschätzten Planungskompetenz festzustellen. Hier schätzen die naturwissenschaftlich orientierten Schüler ihre Planungskompetenz etwas höher ein als die nicht naturwissenschaftlich orientierten ($t(163)=2,57$; $p \leq 0,05$; $\omega^2=0,050$). Diese mäßigen Unterschiede entsprechen den Beobachtungen im Pretest. Statistisch bedeutsame Unterschiede zwischen den Schülertypen sind in allen Versuchsgruppen zu belegen (Genetischer Fingerabdruck: ($F(2,209)=4,86$; $p \leq 0,01$; $\eta^2=0,044$); LDH: ($F(2,171)=19,85$; $p \leq 0,001$; $\eta^2=0,188$); Einzeller: ($F(2,61)=3,80$; $p \leq 0,05$; $\eta^2=0,111$). Mittels Post-hoc-Tests sind diese im Laborkurs Genetischer Fingerabdruck auf eine signifikant höhere Kompetenzeinschätzung der

Allrounder gegenüber den Theoretikern (Scheffé; $p \leq 0,05$) und im Kurs LDH auf eine bedeutend höhere Kompetenzeinschätzung der Allrounder gegenüber den Praktikern und Theoretikern zurückzuführen (Scheffé jeweils $p \leq 0,001$). In der Versuchsgruppe Einzeller schätzen die Allrounder ihre Planungskompetenz höher ein als die Praktiker (Scheffé; $p \leq 0,05$). Die im Labor empfundene Planungskompetenz der Praktiker unterscheidet sich somit nicht nachweisbar von derjenigen der Theoretiker, was in etwa mit den Beobachtungen im Pretest übereinstimmt.

Einen Eindruck, wie stark die dispositionale Selbstwirksamkeitserwartung mit der tatsächlich im Labor empfundenen Selbstwirksamkeit übereinstimmt, gibt Tabelle 126. Hier ist die unmittelbar im Anschluss an den Laborbesuch erhobene Selbstwirksamkeitserwartung in der Subskala Planung korreliert mit den in Pre- und Follow-up-Test erhobenen Einschätzungen.

Tabelle 126: Korrelationen (r), Signifikanz und Stichprobengröße (N) der postFSWEx.V1-Subskala Planung mit der entsprechenden FSWEx.V1-Skala in Pre und Follow-up-Test für die drei Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller

			Genetischer Fingerabdruck		LDH		Einzeller	
			FSWEx.V1 – Planung (Pretest)	FSWEx.V1 – Planung (Follow-up-Test)	FSWEx.V1 – Planung (Pretest)	FSWEx.V1 – Planung (Follow-up-Test)	FSWEx.V1 – Planung (Pretest)	FSWEx.V1 – Planung (Follow-up-Test)
Genetischer Fingerabdruck	postFSWEx.V1 – Planung (Posttest)	r Signifikanz N	0,280 $p \leq 0,001$ 164	0,267 $p \leq 0,001$ 164	---	---	---	---
	FSWEx.V1 – Planung (Pretest)	r Signifikanz N	1	0,625 $p \leq 0,001$ 164	---	---	---	---
LDH	postFSWEx.V1 – Planung (Posttest)	r Signifikanz N	---	---	0,477 $p \leq 0,001$ 115	0,424 $p \leq 0,001$ 115	---	---
	FSWEx.V1 – Planung (Pretest)	r Signifikanz N	---	---	1	0,554 $p \leq 0,001$ 115	---	---
Einzeller	postFSWEx.V1 – Planung (Posttest)	r Signifikanz N	---	---	---	---	0,331 $p \leq 0,01$ 64	Keine Daten vorhanden

Ersichtlich ist, dass die Pre- und Follow-up-Test-Werte signifikant positiv mit der im Labor empfundenen Selbstwirksamkeit korrelieren. Insbesondere in den Kursen Genetischer Fingerabdruck und Einzeller ist diese Korrelation mit $r < 0,35$ allerdings nicht besonders hoch. Es ist also nicht so, dass alle Schüler, die vor dem Laborbesuch ihre Planungskompetenz hoch einschätzen auch in der konkreten Laborsituation denken, dass sie das durchgeführte Experiment auch gut hätten selbst

planen können. Die Korrelationen zwischen Pre- und Follow-up-Test fallen dagegen mit $r > 0,55$ deutlich höher aus. Schüler, die ihre Planungskompetenz im Pretest hoch einschätzen, weisen somit auch im Follow-up-Test häufiger noch die höheren Einschätzungen auf. Diese Beobachtung unterstützt den Eindruck der im vorherigen Abschnitt festgestellten zeitlichen Stabilität dieser Selbstwirksamkeitssubskala.

3.2.2.3.4.3 FSWEx.V1 – Subskala Durchführung

Vor dem Laborbesuch schätzen die beiden Versuchsgruppen und die Kontrollgruppe ihre Selbstwirksamkeit, Experimente erfolgreich durchführen zu können, nur geringfügig unterschiedlich ein ($F(2,403)=2,15$; $p=0,118$). Während sich die Selbstwirksamkeitserwartung der Kontrollgruppe nicht ändert, ist für beide Versuchsgruppen im Follow-up-Test eine geringfügige Erhöhung gegenüber dem Pretest zu beobachten (Tab. 127 & Abb. 75). Durch eine zweifaktorielle ANOVA mit Messwiederholung ist allerdings keine signifikante Steigerung der Selbstwirksamkeitserwartung in den Versuchsgruppen gegenüber der Kontrollgruppe nachzuweisen (Interaktion Zeit x Treatment: $F(2,403)=1,50$; $p=0,224$).

Tabelle 127: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie der Stichprobengröße (*N*) der Treatmentgruppen in der FSWEx.V1-Subskala Durchführung im Pre- und Follow-up-Test.

Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Genetischer Fingerabdruck	2,59	0,76	171
	LDH	2,78	0,84	130
	Kontrollgruppe	2,66	0,72	105
Follow-up-Test	Genetischer Fingerabdruck	2,77	0,77	171
	Versuchsgruppe LDH	2,94	0,80	130
	Kontrollgruppe	2,69	0,82	105

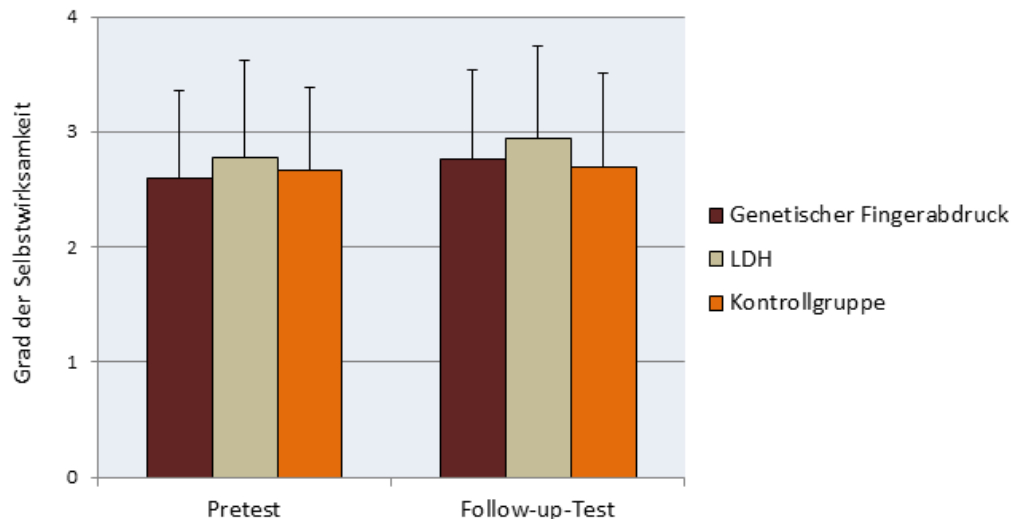


Abbildung 75: Mittelwerte der FSWEx.V1-Subskala Durchführung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4).

Innerhalb der Versuchsgruppen liegt die Selbstwirksamkeitserwartung, Experimente erfolgreich durchführen zu können, für weibliche und männliche Kursteilnehmer auf gleichem Niveau. Für beide Geschlechter ist in beiden Versuchsgruppen gleichermaßen die bereits beschriebene leichte Erhöhung der Selbstwirksamkeit festzustellen. Gleiches gilt für Schüler mit und ohne Migrationshintergrund. Naturwissenschaftlich orientierte Kursteilnehmer zeichnen sich dagegen bereits im Pretest beider Versuchsgruppen durch eine signifikant höher eingeschätzte Durchführungskompetenz aus (Genetischer Fingerabdruck: $t(153)=2,65$; $p\leq 0,01$; $\omega^2=0,037$; LDH: $t(112)=2,75$; $p\leq 0,01$; $\omega^2=0,054$). In beiden Versuchsgruppen wird die Selbstwirksamkeitserwartung der naturwissenschaftlich und nicht naturwissenschaftlich Orientierten leicht gesteigert. In beiden Gruppen fällt diese Steigerung bei den naturwissenschaftlich orientierten signifikant höher aus als bei den nicht naturwissenschaftlich orientierten Schülern (Haupteffekt Berufsorientierung: Genetischer Fingerabdruck: $F(1,152)=4,43$; $p\leq 0,05$; $\eta^2=0,028$; LDH: $F(1,111)=8,04$; $p\leq 0,01$; $\eta^2=0,068$). Die drei Schülertypen unterscheiden sich zum Zeitpunkt des Pretests sowohl in der Versuchsgruppe Genetischer Fingerabdruck ($F(2,165)=21,76$; $p\leq 0,001$; $\eta^2=0,209$), als auch in der Gruppe LDH ($F(2,124)=7,84$; $p\leq 0,001$; $\eta^2=0,112$) signifikant. In beiden Versuchsgruppen ist dies auf eine mindestens sehr signifikant höhere Selbstwirksamkeitserwartung der Allrounder gegenüber den Praktikern und Theoretikern zurückzuführen (Scheffé jeweils mindestens $p\leq 0,01$). Wie Tabelle 128 und Abbildung 76 zu entnehmen ist, schätzen die Praktiker ihre Durchführungskompetenz vor dem Laborkurs in der Versuchsgruppe Genetischer Fingerabdruck höher ein als die Theoretiker. In der Versuchsgruppe LDH liegen sie im Pretest auf ähnlichem Niveau mit den Theoretikern.

Zum Follow-up-Test hin ändert sich die Selbstwirksamkeit, Experimente erfolgreich durchzuführen, bei den Allroundern und Praktikern kaum. Bei den Allroundern steigt sie geringfügig an und bei den Praktikern sinkt sie ein wenig. Die Selbstwirksamkeitserwartung der Theoretiker steigt dagegen

deutlich an. Mittels ANCOVA (Pretest als Kovariate) wird in beiden Versuchsgruppen ein signifikanter Haupteffekt des Schülertypus auf die im Follow-up-Test gefundenen Unterschiede bestätigt (Genetischer Fingerabdruck: $F(2,164)=4,43$; $p \leq 0,05$; $\eta^2=0,051$; LDH: $F(2,123)=7,00$; $p \leq 0,001$; $\eta^2=0,102$). Dieser kann in beiden Versuchsgruppen mittels Post-hoc-Test nach Bonferroni auf signifikante Unterschiede in der Änderung der Selbstwirksamkeitserwartung zwischen den Theoretikern und Praktikern sowie den Allroundern und Praktikern zurückgeführt werden (Bonferroni mindestens jeweils $p \leq 0,05$). Der Anstieg der Selbstwirksamkeitserwartung bei den Allroundern und insbesondere bei den Theoretikern unterscheidet diese Gruppen somit signifikant von den Praktikern (Tab. 128 & Abb.76-77).

Tabelle 128: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie die Stichprobengrößen (*N*) der Versuchsgruppen Genetischer Fingerabdruck (oben) und LDH (unten) in der FSWEx.V1-Subskala Durchführung im Pre- und Follow-up-Test (vgl. Abb. 76 & 77)

Genetischer Fingerabdruck	Messzeitpunkt	Schülertypus	<i>M</i>	<i>SD</i>	<i>N</i>
	Pretest	Allrounder	2,91	0,71	83
		Praktiker	2,45	0,66	31
		Theoretiker	2,14	0,64	54
	Follow-up-Test	Allrounder	3,02	0,66	83
		Praktiker	2,39	0,85	31
		Theoretiker	2,56	0,73	54
LDH	Messzeitpunkt	Schülertypus	<i>M</i>	<i>SD</i>	<i>N</i>
	Pretest	Allrounder	3,13	0,59	54
		Praktiker	2,65	0,90	47
		Theoretiker	2,54	0,69	26
	Follow-up-Test	Allrounder	3,28	0,63	54
		Praktiker	2,57	0,86	47
		Theoretiker	2,96	0,76	26

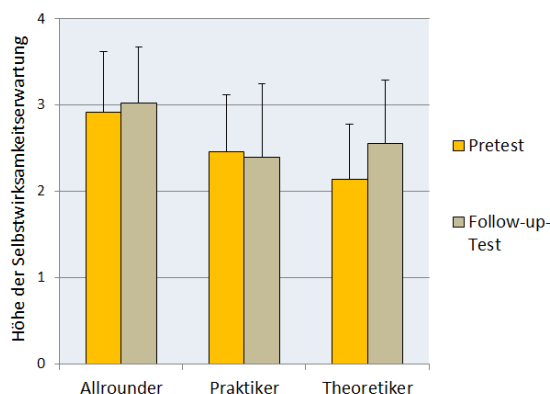


Abbildung 76: Mittelwerte der FSWEx.V1-Subskala Durchführung der Allrounder, Praktiker und Theoretiker in Pre- und Follow-up-Test der Versuchsgruppe Genetischer Fingerabdruck inklusive Standardabweichungen (Skalenmaximum=4)

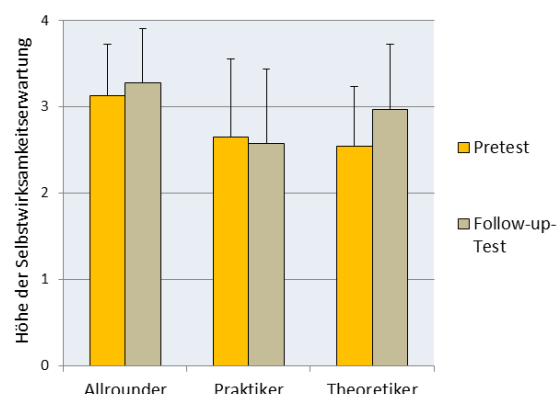


Abbildung 77: Mittelwerte der FSWEx.V1-Subskala Durchführung der Allrounder, Praktiker und Theoretiker in Pre- und Follow-up-Test der Versuchsgruppe LDH inklusive Standardabweichungen (Skalenmaximum=4)

3.2.2.3.4.4 postFSWEx.V1 – Subskala Durchführung

Wie Tabelle 129 und Abbildung 78 verdeutlichen, schätzen die Schüler ihre Durchführungskompetenz auch unmittelbar auf die BeLL Bio-Laborkurse bezogen sehr hoch ein. Besonders auffällige Unterschiede sind zwischen den Kursen nicht zu beobachten. Dennoch lässt sich im Laborkurs Einzeller eine gegenüber dem Laborkurs LDH signifikant niedrigere wahrgenommene Durchführungskompetenz feststellen ($F(2,466)=4,79$; $p\leq 0,01$; $\eta^2=0,020$).

Tabelle 129: Mittelwerte (M) und Standardabweichungen (SD) sowie die Stichprobengrößen (N) der Versuchsgruppen in der postFSWEx.V1-Subskala Durchführung im Posttest

Messzeitpunkt	Treatmentgruppe	M	SD	N
Posttest	Genetischer Fingerabdruck	3,11	0,67	216
	LDH	3,27	0,62	184
	Einzeller	3,01	0,73	69

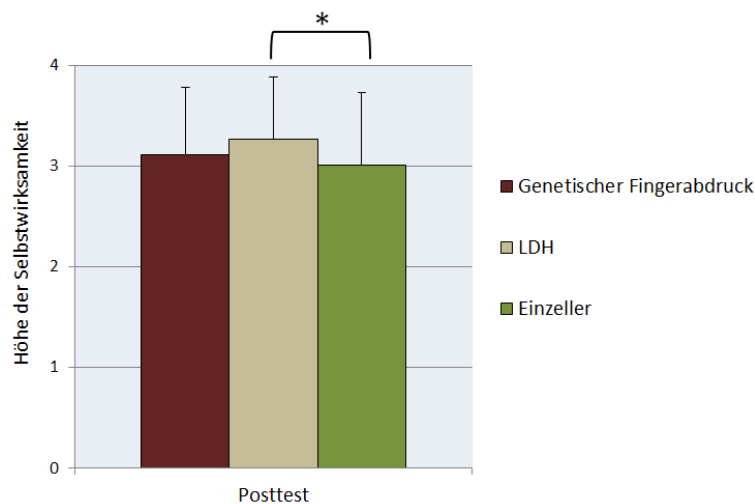


Abbildung 78: Mittelwerte der postFSWEx.V1-Subskala Durchführung der Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller im Posttest inklusive Standardabweichungen (Skalenmaximum=4; *: $p\leq 0,05$)

Es folgen die Subgruppenvergleiche innerhalb der Versuchsgruppen. Bezüglich des Migrationshintergrundes und der beruflichen Orientierung können für den Einzellerkurs keine Aussagen getroffen werden, da keine entsprechenden Daten in diesem Kurs erhoben wurden. Innerhalb der Versuchsgruppen können in keinem der Kurse signifikante, vom Geschlecht oder Migrationshintergrund abhängige Unterschiede festgestellt werden. In der Versuchsgruppe Einzeller schätzen die naturwissenschaftlich orientierten ihre Durchführungskompetenz tendenziell ($t(147)=1,81$; $p=0,073$) und in der Gruppe LDH signifikant ($t(103)=2,53$; $p\leq 0,05$; $\omega^2=0,049$) höher ein, als die nicht naturwissenschaftlich ausgerichteten Schüler. Signifikante Unterschiede zwischen den Schülertypen sind in jeder Versuchsgruppe festzustellen (Genetischer Fingerabdruck: $F(2,209)=3,80$; $p\leq 0,001$; $\eta^2=0,035$; Einzeller: $F(2,61)=3,48$; $p\leq 0,05$; $\eta^2=0,102$; LDH: $F(2,171)=6,06$; $p\leq 0,01$; $\eta^2=0,066$). Im Kurs Genetischer Fingerabdruck schätzen die Allrounder ihre Durchführungskompetenz signifikant

höher ein, als die Theoretiker (Scheffé; $p \leq 0,05$). Der varianzanalytisch aufgespürte Unterschied in der Gruppe Einzeller lässt sich mittels Post-hoc-Test nur tendenziell nachweisen. Hier zeigen die Allrounder eine etwas höhere Selbsteinschätzung als die Praktiker (Scheffé; $p = 0,60$). In der Versuchsgruppe Einzeller ist der Unterschied der Allrounder zu den Praktikern signifikant (Scheffé; $p \leq 0,01$). Theoretiker und Praktiker schätzen ihre Durchführungskompetenz somit nicht nachweisbar unterschiedlich ein.

In Tabelle 130 wird die im Laborkurs wahrgenommene Durchführungskompetenz korreliert mit den Selbstwirksamkeitserwartungen in Pre- und Posttest. Im Laborkurs Genetischer Fingerabdruck ähneln die Korrelationen denen, welche bereits für die Subskala Planung festgestellt wurden. Während die Posttestergebnisse nur relativ gering mit den Einschätzungen in Pre- und Follow-up-Test korrelieren, ist ein relativ hoher Zusammenhang zwischen den Pre- und Follow-up-Werten zu finden. Somit schätzen Schüler ihre Durchführungskompetenz unmittelbar im Anschluss an den Laborkurs Genetischer Fingerabdruck häufig anders ein als dies ihrer generellen Selbstwirksamkeitserwartung entspricht. Gleiches gilt für die Versuchsgruppe Einzeller. Etwas anders sehen die Zusammenhänge allerdings in der Versuchsgruppe LDH aus. Hier besteht zwischen den Einschätzungen in Pre- und Posttest mit $r(115) = 0,185$; $p \leq 0,05$ ebenfalls nur ein geringer Zusammenhang.

Tabelle 130: Pearson-Korrelationen (r), Signifikanz und Stichprobengröße (N) der postFSWEx.V1-Subskala Durchführung mit den FSWEx.V1-Skalen in Pre und Follow-up-Test für die drei Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller

			Genetischer Fingerabdruck		LDH		Einzeller	
			FSWEx.V1 – Durchführung (Pretest)	FSWEx.V1 Durchführung (Follow-up-Test)	FSWEx.V1 – Durchführung (Pretest)	FSWEx.V1 Durchführung (Follow-up-Test)	FSWEx.V1 – Durchführung (Pretest)	FSWEx.V1 Durchführung (Follow-up-Test)
Genetischer Fingerabdruck	postFSWEx.V1 – Durchführung (Posttest)	r Signifikanz N	0,154 $p \leq 0,01$ 164	0,245 $p \leq 0,01$ 164	---	---	---	---
	FSWEx.V1 – Durchführung (Pretest)	r Signifikanz N	1	0,599 $p \leq 0,001$ 164	---	---	---	---
LDH	postFSWEx.V1 – Durchführung (Posttest)	r Signifikanz N	---	---	0,185 $p \leq 0,05$ 115	0,498 $p \leq 0,001$ 115	---	---
	FSWEx.V1 – Durchführung (Pretest)	r Signifikanz N	---	---	1	0,458 $p \leq 0,001$ 115	---	---
Einzeller	postFSWEx.V1 – Durchführung (Posttest)	r Signifikanz N	---	---	---	---	0,119 $p = 0,350$ 64	Keine Daten vorhanden

Allerdings ist die Korrelation zwischen der Kompetenzeinschätzung im Posttest und im Follow-up-Test $r(115)=0,498$; $p \leq 0,001$ auffallend hoch. Der Zusammenhang zwischen Pre- und Follow-up-Test ist zwar auch noch relativ stark, aber nicht so ausgeprägt wie dies in der Versuchsgruppe Genetischer Fingerabdruck der Fall ist. Schüler, die ihre Durchführungskompetenz im Laborkurs LDH hoch einschätzen, tun dies somit häufig auch noch im Follow-up-Test. In Verbindung mit der Beobachtung, dass sich die Kompetenzeinschätzung vor dem Laborbesuch wenig mit der im Laborbesuch deckt, deutet sich eine zeitliche Veränderung der wahrgenommenen Durchführungskompetenz an.

3.2.2.3.4.5 FSWEx.V1 – Subskala Auswertung

Wie bei den FSWEx-Subskalen Planung und Durchführung werden zunächst die Pre- und Follow-up-Test-Resultate vorgestellt. Zwischen den Versuchsgruppen und der Kontrollgruppe sind im Pretest keine Unterschiede in der Selbstwirksamkeitserwartung, Experimente erfolgreich auswerten zu können, festzustellen ($F(2,397)=0,10$; $p=0,907$). Für die beiden Versuchsgruppen kann eine sehr geringfügige Zunahme der Selbstwirksamkeitserwartung im Follow-up-Test festgestellt werden, welche allerdings gegenüber der Kontrollgruppe nicht signifikant ist (Interaktion Zeit x Treatment: $F(2,397)=0,58$; $p=0,562$). Tabelle 131 und Abbildung 79 stellen die Selbstwirksamkeitserwartung der Schüler in der Subskala Auswertung beider Versuchsgruppen und der Kontrollgruppe dar.

Tabelle 131: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengröße (*N*) der Treatmentgruppen in der FSWEx.V1-Subskala Auswertung im Pre- und Follow-up-Test

Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Genetischer Fingerabdruck	2,34	0,65	170
	LDH	2,32	0,75	127
	Kontrollgruppe	2,31	0,62	103
Follow-up-Test	Genetischer Fingerabdruck	2,45	0,80	170
	LDH	2,42	0,77	127
	Kontrollgruppe	2,34	0,81	103

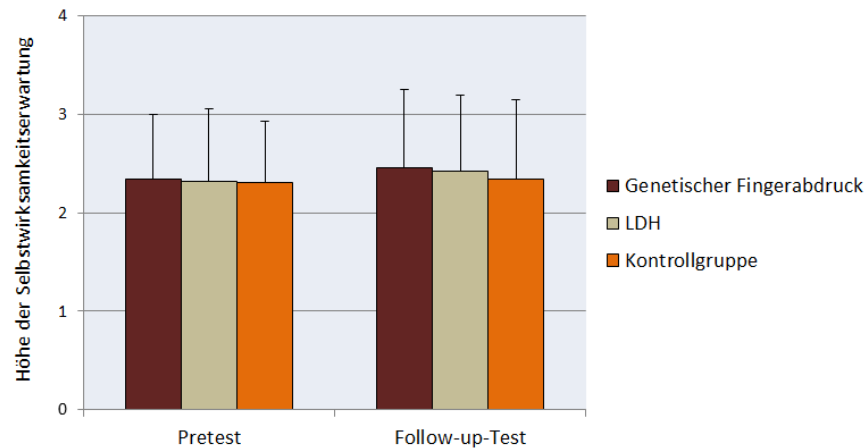


Abbildung 79: Mittelwerte der FSWEx.V1-Subskala Auswertung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

Zum Zeitpunkt des Pretests schätzen in beiden Versuchsgruppen die männlichen Kursteilnehmer ihre Auswertungskompetenz geringfügig – aber nicht signifikant – höher ein als die weiblichen Teilnehmer. In der Versuchsgruppe Genetischer Fingerabdruck geben beide Geschlechter zum Zeitpunkt des Follow-up-Tests eine geringfügig höhere Selbsteinschätzung an als im Pretest. Bei den weiblichen Kursteilnehmern in der Versuchsgruppe LDH ist dagegen keine Zunahme der Selbstwirksamkeitserwartung zu beobachten. Da bei den männlichen Schülern eine Erhöhung zu verzeichnen ist, unterscheiden sich diese in einer zweifaktoriellen ANOVA mit Messwiederholung tendenziell von den weiblichen Kursteilnehmern (Interaktion Zeit x Geschlecht: $F(1,125)=3,31$; $p=0,071$). Zum Zeitpunkt des Pretests schätzen Schüler mit und ohne Migrationshintergrund in beiden Versuchsgruppen ihre Fähigkeit, Experimente auszuwerten, gleich ein. Statistisch bedeutsame Veränderungen zum Follow-up-Test sind nicht zu belegen. Wie dies bereits für die beiden anderen Subskalen des FSWEx.V1 gezeigt wurde, so geben auch in der Subskala Auswertung die Schüler mit naturwissenschaftlicher beruflicher Orientierung im Pretest eine signifikant höhere Selbstwirksamkeitserwartung an als die nicht naturwissenschaftlich orientierten (Genetischer Fingerabdruck: $t(152)=2,98$; $p\leq 0,01$; $\omega^2=0,049$; LDH: $t(109)=3,26$; $p\leq 0,001$; $\omega^2=0,080$). Bedeutsame Unterschiede zwischen den beiden Gruppen bezüglich der im Follow-up-Test beobachteten Erhöhung der Selbstwirksamkeitserwartung sind nicht zu nachzuweisen. Auch zwischen den Schülertypen können im Pretest wieder deutliche Unterschiede nachgewiesen werden (Genetischer Fingerabdruck: $F(2,165)=21,83$; $p\leq 0,001$; $\eta^2=0,209$; LDH: $F(2,124)=12,70$; $p\leq 0,001$; $\eta^2=0,170$). Wieder äußern die Allrounder die höchsten Selbstwirksamkeitserwartungen. Allerdings schätzen die Theoretiker ihre Auswertungskompetenz ähnlich hoch ein, mit einigem Abstand gefolgt von den Praktikern. Post-hoc-Tests bestätigen in beiden Versuchsgruppen eine signifikant höhere Selbstwirksamkeitserwartung der Allrounder und der Theoretiker gegenüber den Praktikern (Scheffé jeweils mindestens $p\leq 0,05$). Im Falle der Versuchsgruppe Genetischer Fingerabdruck ist zudem die

Selbstwirksamkeitserwartung der Allrounder signifikant höher als die der Theoretiker (Scheffé $p \leq 0,01$). Im Follow-up-Test schätzen sich alle Schüler, unabhängig von Versuchsgruppe und Schülertypus, geringfügig kompetenter ein als im Pretest.

3.2.2.3.4.6 postFSWEx.V1 – Subskala Auswertung

Tabelle 132 und Abbildung 80 stellen für jede der Versuchsgruppen dar, wie die Schüler im Laborkurs ihre Kompetenz einschätzen, Experimente auszuwerten. Bedeutsame Unterschiede zwischen den drei Laborkursen sind nicht festzustellen.

Tabelle 132: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der Versuchsgruppen in der postFSWEx.V1-Subskala Auswertung im Posttest

Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Posttest	Genetischer Fingerabdruck	2,62	0,85	216
	LDH	2,62	0,78	184
	Einzeller	2,47	0,78	69

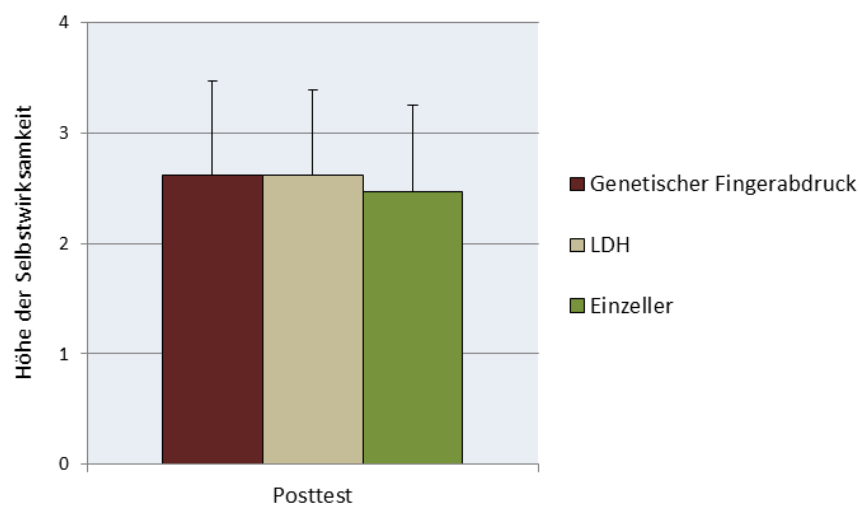


Abbildung 80: Mittelwerte der FSWEx.V1-Subskala Auswertung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

Auch für die Subskala Auswertung erfolgt innerhalb der Versuchsgruppen eine differenzierte Subgruppenanalyse. Diese wird im Kurs Einzeller nur für die Faktoren Geschlecht und Schülertypus vorgenommen, da Daten zur Bestimmung des Migrationshintergrundes und der beruflichen Orientierung in dieser Versuchsgruppe nicht erhoben wurden. Vom Geschlecht oder Migrationshintergrund abhängige Unterschiede sind nicht festzustellen. Sowohl in der Versuchsgruppe Genetischer Fingerabdruck ($t(147)=1,99$; $p \leq 0,05$; $\omega^2=0,019$), als auch in der Gruppe LDH ($t(103)=3,37$; $p \leq 0,001$; $\omega^2=0,090$) schätzen die naturwissenschaftlich orientierten Schüler ihre Auswertungskompetenz signifikant höher ein als die nicht naturwissenschaftlich ausgerichteten.

Signifikante Unterschiede zwischen den Schülertypen sind in den Versuchsgruppen Genetischer Fingerabdruck ($F(2,209)=9,78$; $p\leq 0,001$; $\eta^2=0,086$) und LDH ($F(2,171)=12,24$; $p\leq 0,001$; $\eta^2=0,125$) zu beobachten. Post-hoc-Tests belegen im Kurs Genetischer Fingerabdruck eine signifikant höhere Selbsteinschätzung der Allrounder und der Theoretiker gegenüber den Praktikern (Scheffé jeweils mindestens $p\leq 0,05$). Auch in der Versuchsgruppe LDH schätzen die Praktiker ihre Auswertungskompetenz deutlich niedriger ein als die Allrounder (Scheffé; $p\leq 0,05$) und die Theoretiker (Scheffé; $p=0,104$). Im Laborkurs Einzeller unterscheiden sich die Schülertypen in diesem Punkt nicht.

Wie bereits für die Subskalen Planung und Durchführung vorgestellt, so finden sich in der nachfolgenden Tabelle 133 die Korrelationen der postFSWEx.V1-Subskala Auswertung mit den FSWEx.V1 –Einschätzungen in Pre- und Follow-up-Test. Diese unterscheiden sich relativ deutlich von den Korrelationen der beiden anderen Subskalen. Auch für die Subskala Auswertung sind zwar die höchsten Korrelationen zwischen Pre- und Follow-up-Test zu beobachten, allerdings korrelieren auch die Posttest-Einschätzungen relativ deutlich mit den Pre-, insbesondere aber mit den Follow-up-Test-Daten.

Tabelle 133: Pearson-Korrelationen (r), Signifikanzen sowie Stichprobengrößen (N) der postFSWEx.V1-Subskala Auswertung mit den FSWEx.V1-Skalen in Pre und Follow-up-Test für die drei Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller

			Genetischer Fingerabdruck		LDH		Einzeller	
			FSWEx.V1 – Auswertung (Pretest)	FSWEx.V1 Auswertung (Follow-up-Test)	FSWEx.V1 – Auswertung (Pretest)	FSWEx.V1 Auswertung (Follow-up-Test)	FSWEx.V1 – Auswertung (Pretest)	FSWEx.V1 Auswertung (Follow-up-Test)
Genetischer Fingerabdruck	postFSWEx.V1 – Auswertung (Posttest)	r Signifikanz N	0,478 $p\leq 0,001$ 164	0,483 $p\leq 0,001$ 164	---	---	---	---
	FSWEx.V1 – Auswertung (Pretest)	r Signifikanz N	1	0,657 $p\leq 0,001$ 164	---	---	---	---
LDH	postFSWEx.V1 – Auswertung (Posttest)	r Signifikanz N	---	---	0,355 $p\leq 0,001$ 115	0,520 $p\leq 0,001$ 115	---	---
	FSWEx.V1 – Auswertung (Pretest)	r Signifikanz N	---	---	1	0,591 $p\leq 0,001$ 115	---	---
Einzeller	postFSWEx.V1 – Auswertung (Posttest)	r Signifikanz N	---	---	---	---	0,316 $p\leq 0,05$ 64	Keine Daten vorhanden

Die Selbsteinschätzung der Auswertungskompetenz der Schüler stimmt also offensichtlich eher mit ihrer im Labor wahrgenommenen Auswertungskompetenz überein als dies bei den Subskalen Planung und Durchführung der Fall ist. Wie die relativ hohen Korrelationen zwischen Post- und Follow-up-Test zeigen, stimmt die Einschätzung einige Wochen nach dem Laborbesuch noch deutlicher mit der Selbsteinschätzung im Labor überein als dies vor dem Laborbesuch der Fall ist. Dies deutet – wie bereits im LDH-Kurs für die Subskala Durchführung zu beobachten war – auf eine zeitlich qualitative Veränderung der Selbsteinschätzung hin.

Wie im vorangegangenen Abschnitt gezeigt wurde, lassen sich in vielen der über mehrere Messzeitpunkte erhobenen Daten bedeutsame laborbesuchsbedingte Veränderungen nachweisen. Im folgenden Punkt soll der Einfluss des einmaligen Laborbesuches auf das Berufsinteresse vorgestellt werden.

3.2.2.4 Berufsinteresse

Im Follow-up-Test der Versuchsgruppen Genetischer Fingerabdruck und LDH wurden die Schüler gebeten zu beurteilen, ob durch den BeLL Bio-Besuch ihr Interesse für naturwissenschaftliche Berufe geweckt worden sei.

Tabelle 134: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der Versuchsgruppen in der Skala Berufsinteresse

Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Genetischer Fingerabdruck	1,10	0,83	171
LDH	1,26	0,98	130

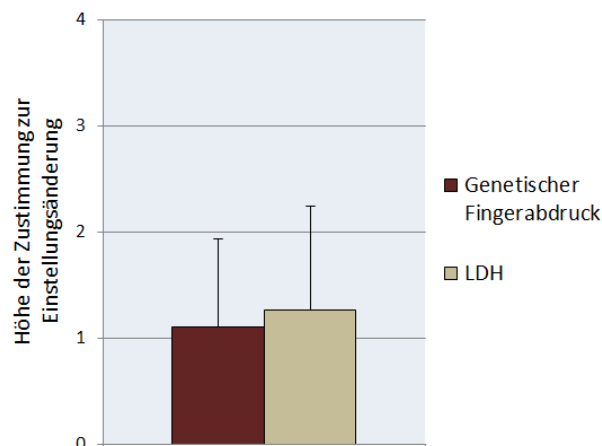


Abbildung 81: Mittelwerte und Standardabweichung der Einschätzung, wie sehr der BeLL Bio Besuch das Schülerinteresse für naturwissenschaftliche Berufe geweckt hat (Skalenmaximum=4)

Tabelle 134 und Abbildung 81 geben einen Überblick, wie sehr die Schüler dieser Überlegung zustimmen. Wie die Abbildung zeigt, ist die durch den einmaligen Laborbesuch bedingte Einstellungsänderung gering, aber vorhanden. Bedeutsame Unterschiede zwischen beiden Versuchsgruppen sind nicht zu belegen ($t(299)=1,54$; $p=0,126$).

Unterschiede im Grad der Zustimmung in Abhängigkeit von Geschlecht und Migrationshintergrund sind in keiner der Versuchsgruppen festzustellen. Allerdings zeichnen sich die Schüler mit naturwissenschaftlicher beruflicher Orientierung durch eine signifikant höhere Zustimmung aus als diejenigen ohne naturwissenschaftliches Berufsziel (Genetischer Fingerabdruck: $t(52)=3,00$; $p\leq 0,01$; $\omega^2=0,050$; LDH: $t(52)=4,60$; $p\leq 0,001$; $\omega^2=0,150$, vgl. Tab. 135 & Abb. 82 & 83).

Tabelle 135: Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der Versuchsgruppen in der Skala Berufsinteresse unter Berücksichtigung der Berufsorientierung

Versuchsgruppe	Berufsorientierung	<i>M</i>	<i>SD</i>	<i>N</i>
Genetischer Fingerabdruck	Naturwissenschaftlich	1,47	1,01	40
	Nicht naturwissenschaftlich	0,95	0,68	115
LDH	Naturwissenschaftlich	1,92	0,91	33
	Nicht naturwissenschaftlich	1,03	0,95	81

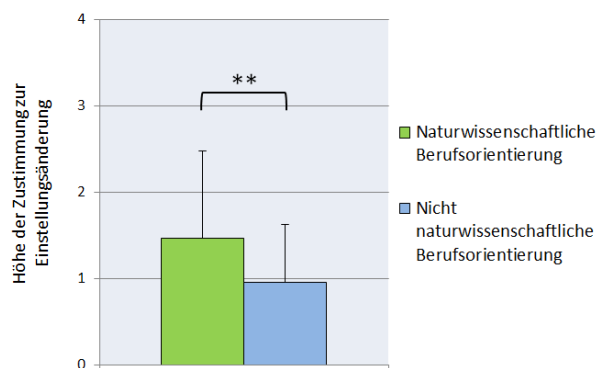


Abbildung 82: Mittelwerte und Standardabweichung der Einschätzung, wie sehr der BeLL Bio-Kurs Genetischer Fingerabdruck das Interesse für naturwissenschaftliche Berufe geweckt hat (Skalenmaximum=4)

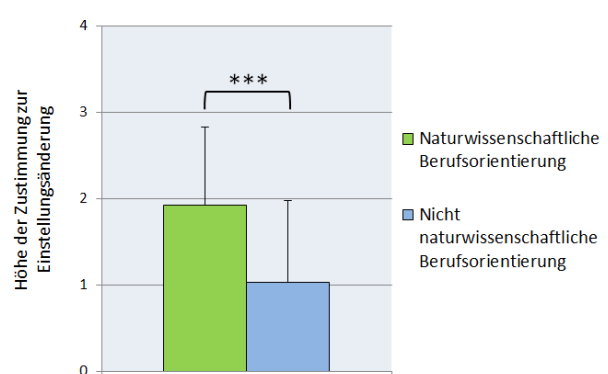


Abbildung 83: Mittelwerte und Standardabweichung der Einschätzung, wie sehr der BeLL Bio-Kurs LDH das Interesse für naturwissenschaftliche Berufe geweckt hat (Skalenmaximum=4)

Auch zwischen den Schülertypen sind bedeutsame Unterschiede zu finden (Genetischer Fingerabdruck: $F(2,165)=7,32$; $p\leq 0,001$; $\eta^2=0,082$; LDH: $F(2,124)=13,51$; $p\leq 0,001$; $\eta^2=0,179$). In beiden Versuchsgruppen geben die Allrounder die höchste laborbesuchsbedingte Interessenzunahme an. Die Erhöhung des Berufsinteresses ist bei den Theoretikern im LDH-Kurs aber annähernd gleich hoch. Hier geben die Praktiker die deutlich geringste Einschätzung ab. Dies verdeutlicht auch der signifikante Unterschied der Allrounder (Scheffé; $p\leq 0,001$) und der Theoretiker

(Scheffé; $p \leq 0,001$) zu den Praktikern. Im Laborkurs Genetischer Fingerabdruck bekunden die Praktiker eine ähnlich hohe Interessenzunahme wie die Theoretiker. Beide Gruppen unterscheiden sich signifikant von den Allroundern (Scheffé; $p \leq 0,001$).

Bisher wurden in einzelnen Abschnitten die Ergebnisse der erhobenen Personen- und Laborvariablen sowie der längsschnittlichen Untersuchungen berichtet. Innerhalb der Personenvariablen wurden zudem bereits Korrelationen aufgeführt, welche zwar das Vorliegen oder Nicht-Vorliegen eines Zusammenhanges offenlegen, aber keine Aussagen über die kausale Richtung eines Zusammenhangs zulassen. Das folgende lineare Strukturgleichungsmodell stellt viele der bisher vorgestellten Variablen in ein kausales Beziehungsnetz.

3.2.2.5 Modell zur affektiven Wirkung des BeLL Bio-Besuches

An dieser Stelle wird das bereits postulierte Modell affektiver Einflussgrößen am Lernort BeLL Bio entsprechend des beschriebenen Vorgehens überprüft. Das Modell wurde in Anlehnung an das Strukturgleichungsmodell über die Interessenförderung durch Schülerlabore (Pawek 2009) entworfen. Nähere Erläuterungen zur Strukturgleichungsmodellierung finden sich unter Punkt 2.2.5.10. Einbezogen in das hier berichtete Strukturgleichungsmodell (Abb. 84) werden als latente Personenvariablen das Fach- und Sachinteresse sowie das biologische Fähigkeitsselbstkonzept (FSK) und die experimentbezogene Selbstwirksamkeit (SelbstExp) vor dem Laborbesuch. Als manifeste Variablen des Fach- und Sachinteresses sowie des Fähigkeitsselbstkonzeptes dienen die entsprechenden Items der im Abschnitt 2.2.3.2.3.1 beschriebenen Messinstrumente. Die manifesten Variablen für die experimentbezogene Selbstwirksamkeit bilden die Summenscores der drei FSWEx.V1-Subskalen Planung (Plan), Durchführung (Durch) und Auswertung (Ausw). Von den Personenvariablen ausgehende, fett gedruckte blaue Pfeile kennzeichnen den Einfluss dieser Variablen auf den Laborbesuch. Als manifeste, den Laborbesuch repräsentierende Variablen finden die Summenscores der Messinstrumente für die Authentizität, den Alltagsbezug, die Transparenz der Versuchsanleitung sowie die Zusammenarbeit mit den Mitschülern Verwendung. Für die manifesten Variablen Atmosphäre und Betreuer wurden aufgrund der mangelhaften Skalenreliabilität die Einzelitems poa4 und poa44 verwendet. Die Variable Offenheit wurde nicht ins Modell aufgenommen, da sich dieses Konstrukt in anderen Studien als nur mäßig relevanter Einflussfaktor am Lernort Schülerlabor herausstellte (vgl. Engeln 2004). Auf die Aufnahme der Variable Druck in das Modell wurde verzichtet, um Verzerrungen durch einen möglichen Bodeneffekt in der Skala zu vermeiden. Vom Laborbesuch ausgehende Effekte auf die experimentbezogene Selbstwirksamkeit (SelbstExp), das aktuelle Interesse (AktInteresse) und das biologische Fähigkeitsselbstkonzept (FSK)

zum Posttestzeitpunkt (post) werden genau wie die Einflüsse auf die Follow-up-Test-Daten (fol) durch fett gedruckte rote Pfeile symbolisiert. Die latente Variable des aktuellen Interesses wird in Anlehnung an Pawek (2009) jeweils durch die Summenscores der drei Interessenkomponenten, also das emotionale (Emo), epistemische (Epi) und das wertebezogene Interesse (Wert) repräsentiert. Als Indikatoren für die latente Variable postSelbstExp finden die Summenscores der Subskalen des postFSWEx Verwendung. Für die latente Variable folSelbstExp werden die Summenscores der Subskalen des im Follow-up-Test eingesetzten FSWEx.V1 als manifeste Variablen eingesetzt. Indikatoren des Messinstrumentes Berufsinteresse sind die Items des entsprechenden Messinstrumentes. Die Position des Berufsinteresses im Modell resultiert daraus, dass dieses Messinstrument erst im Follow-up-Test eingesetzt wurde. Auf die Berücksichtigung des Fachwissens musste bedingt durch die gemeinsame Auswertung der Schüler aus den Treatmentgruppen Genetischer Fingerabdruck und LDH verzichtet werden. Als Pfadkoeffizienten sind die standardisierten β -Gewichte angegeben. β -Gewichte mit einem Betrag $>0,2$ können in Anlehnung an Chin (1998) als bedeutungsvoll beurteilt werden. In der Beschreibung der Modellergebnisse werden als β die direkten Effekte und als β_{total} totale Effekte benannt.

Die Fit-Indices deuten einen akzeptablen Modell-Fit an ($\chi^2/\text{d.f.}=2,269$; RMSEA=0,065; IFI=0,884, CFI=0,882). Lediglich IFI und CFI liegen knapp unter dem in der Literatur geforderten Wert von $\geq 0,90$ (vgl. Abschnitt 2.2.5.10, Tab.74). Entgegen der Modellannahmen erweisen sich die direkten Effekte des dispositionalen, also vor dem Laborbesuch bestehenden Fähigkeitsselbstkonzeptes ($\beta=-0,12$; $p=0,343$), der experimentbezogenen Selbstwirksamkeitserwartung ($\beta=0,17$; $p=0,121$) sowie des Sachinteresses ($\beta=0,10$; $p=0,245$) auf die Laborwahrnehmung als nicht bedeutsam. Lediglich das

Fachinteresse übt einen bedeutsamen ($\beta=0,30$; $p\leq 0,01$) Einfluss auf die Laborwahrnehmung aus. Schüler mit höherem dispositionalem Fachinteresse empfinden im Labor einen höheren Alltagsbezug, höhere Authentizität, eine positivere Atmosphäre und schätzen die Transparenz der Versuchsanleitung höher ein. Weiterhin messen sie den Gesprächen mit dem Betreuer höhere Bedeutung zu und schätzen ihre Zusammenarbeit mit ihren Mitschülern als besser ein. Allerdings ist die Größe des Effektes als moderat zu beurteilen, so dass die persönliche Disposition der Schüler für die Laborwahrnehmung eine eher untergeordnete Rolle spielt. Die postulierten Einflüsse der dispositionalen Variablen untereinander (preSelbstExp \rightarrow Sachinteresse, $\beta=0,30$; $p\leq 0,001$; Sachinteresse \rightarrow Fachinteresse, $\beta=0,30$; $p\leq 0,001$) sind zwar signifikant, bis auf den deutlichen Effekt des Fähigkeitsselbstkonzeptes auf das Fachinteresse an Biologie ($\beta=0,66$; $p\leq 0,001$) allerdings nicht allzu hoch.

Der Effekt der Laborwahrnehmung auf das aktuelle Interesse ($\beta=0,90$; $p\leq 0,001$) unmittelbar nach dem Laborbesuch und die im Labor empfundene experimentbezogene Selbstwirksamkeit ($\beta=0,44$; $p\leq 0,001$) fallen relativ deutlich aus. Im Vergleich dazu ist der Einfluss des Laborbesuches auf das biologische Fähigkeitsselbstkonzept zwar signifikant, aber eher gering ($\beta=0,20$; $p\leq 0,001$). Der deutliche Einfluss des Labors auf das aktuelle Interesse ist insbesondere deshalb von Relevanz, da das dispositionale Fachinteresse keinen ($\beta=0,01$; $p=0,838$) und das vor dem Laborkurs bestehende Sachinteresse nur einen unbedeutenden direkten ($\beta=0,16$; $p\leq 0,01$) Einfluss auf das aktuelle Interesse nimmt. Bedingt durch die geringen, über die Laborwahrnehmung und das Sachinteresse vermittelten indirekten Effekte, fällt der totale Effekt des Sachinteresses auf das aktuelle Interesse mit $\beta_{total}=0,34$ allerdings etwas höher aus. Für das Fachinteresse liegt der totale Effekt auf das aktuelle Interesse bei $\beta_{total}=0,28$. Dennoch ist der Einfluss des Laborbesuches auf das aktuelle Interesse offensichtlich höher als die vor dem Laborkurs bestehende persönliche Disposition. Die im Labor empfundene experimentbezogene Selbstwirksamkeit (postSelbstExp) wird dagegen sehr deutlich von der dispositionalen Selbstwirksamkeitserwartung (preSelbstExp) beeinflusst ($\beta=0,55$; $p\leq 0,001$; $\beta_{total}=0,65$). Der Einfluss des Labors ist hier geringer (BeLL Bio \rightarrow postSelbstExp; $\beta=0,44$; $p\leq 0,001$). Ein besonders starker Einfluss der Disposition (preFSK) ist für das Fähigkeitsselbstkonzept nach dem Laborbesuch (postFSK) festzustellen (preFSK \rightarrow postFSK; $\beta=0,80$; $p\leq 0,001$; $\beta_{total}=0,82$; BeLL Bio \rightarrow postFSK; $\beta=0,20$; $p\leq 0,001$). Bedeutsame nachhaltige, direkte Effekte des Laborbesuches sind nicht festzustellen. Die totalen Effekte fallen allerdings, bedingt durch indirekte Effekte über die im Posttest erhobenen Einstellungen, höher aus (BeLL Bio \rightarrow folAktInteresse; $\beta=-0,16$; $p=0,721$; $\beta_{total}=0,49$; BeLL Bio \rightarrow folFSK; $\beta=0,14$; $p\leq 0,01$; $\beta_{total}=0,20$; BeLL Bio \rightarrow folSelbstExp; $\beta=0,05$; $p=0,512$; $\beta_{total}=0,20$). Sowohl das im Behaltenstest erhobene Interesse, als auch das hier erhobene Fähigkeitsselbstkonzept haben einen signifikanten Effekt auf das durch den Laborbesuch generierte

Interesse an naturwissenschaftlichen Berufen ($\text{folAktInteresse} \rightarrow \text{Berufsinteresse}$; $\beta=0,54$; $p \leq 0,001$; $\text{folFSK} \rightarrow \text{Berufsinteresse}$; $\beta=0,21$, $p \leq 0,01$). Ein Effekt der im Behaltenstest erhobenen experimentbezogenen Selbstwirksamkeit auf das Berufsinteresse ist dagegen nicht zu beobachten ($\beta=0,03$, $p=0,646$). Ein direkter Einfluss des Laborbesuches auf das Berufsinteresse ist nicht festzustellen. Allerdings kann ein moderater indirekter, über das aktuelle Interesse vermittelter Effekt belegt werden ($\beta_{\text{total}}=0,309$). Eine Übersicht aller direkten, indirekten und totalen Effekte des Modells findet sich in Anhang C.

3.2.3 Nebenstudie

Die Nebenstudie soll über die Feststellungen der Hauptstudie hinausgehende Aussagen dazu ermöglichen, ob der einmalige Besuch eines Schülerlabors zu einer kurzfristigen Veränderung der experimentbezogenen Selbstwirksamkeit führt. Dazu wurde im Laborkurs Genetischer Fingerabdruck im Pre-, Post- und Behaltenstest ein identischer FSWEx.V2 eingesetzt, so dass die Ergebnisse hier unmittelbar in eine Längsschnittanalyse über alle drei Messzeitpunkte einfließen können. Nacheinander werden hier die Ergebnisse der drei Subskalen Planung, Durchführung und Auswertung vorgestellt. Für die Schüler der Nebenstudie liegen keine Informationen über Migrationshintergrund, berufliche Orientierung und Schülertypen vor. Somit beschränkt sich die differenziertere Analyse hier auf mögliche Unterschiede zwischen den Geschlechtern.

3.2.3.1 FSWEx.V2 – Subskala Planung

Zum Zeitpunkt des Pretests können keine bedeutsamen Unterschiede in der subjektiv wahrgenommenen Planungskompetenz zwischen Versuchs- und Kontrollgruppe festgestellt werden ($t(133)=0,81$; $p=0,417$). Beide Gruppen schätzen ihre Selbstwirksamkeit Experimente zu planen, relativ gering ein. Eine Varianzanalyse mit Messwiederholung über die drei Messzeitpunkte zeigt keine signifikanten Veränderungen der Versuchsgruppe mit Laborbesuch gegenüber der Kontrollgruppe (Interaktion Zeit x Treatment: $F(2,266)=0,06$; $p=0,942$, Tab. 136 & Abb. 85).

Tabelle 136: Ergebnisse der Nebenstudie -Mittelwerte (M) und Standardabweichungen (SD) sowie Stichprobengrößen (N) der Versuchs- und Kontrollgruppe in der FSWEx.V2-Subskala Planung im Pre-, Post- und Follow-up-Test

Messzeitpunkt	Treatmentgruppe	M	SD	N
Pretest	Versuchsgruppe	1,72	0,75	45
	Kontrollgruppe	1,82	0,63	90
Posttest	Versuchsgruppe	1,79	0,82	45
	Kontrollgruppe	1,91	0,61	90
Follow-up-Test	Versuchsgruppe	1,69	0,76	45
	Kontrollgruppe	1,82	0,64	90

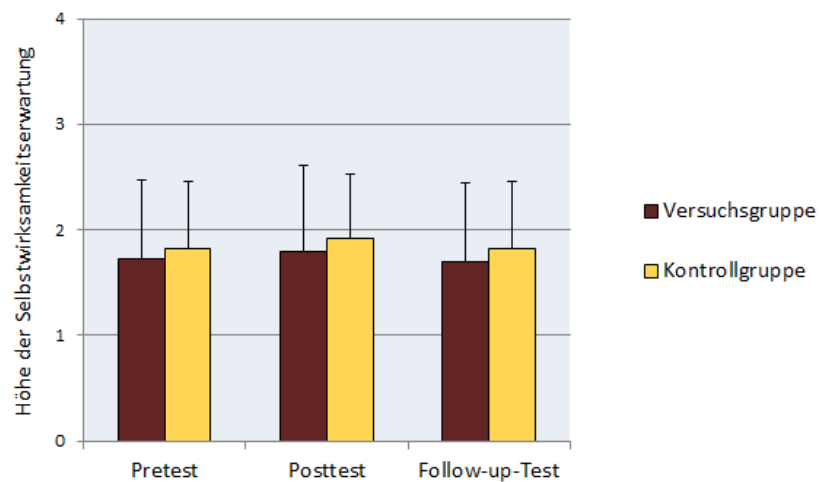


Abbildung 85: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Planung der Versuchs- und Kontrollgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

Ein durch den Laborbesuch bewirkter Zuwachs an subjektiv empfundener Planungskompetenz ist somit auszuschließen. Wie bereits im Rahmen der Hauptstudie festgestellt, bestätigt sich auch hier die langfristige Stabilität der planungsbezogenen Selbstwirksamkeitserwartung (Tabelle 136 & Abbildung 85).

Lohnenswert ist in dieser Subskala der Vergleich der Geschlechter. Bereits im Pretest fällt eine etwas höhere planungsbezogene Selbstwirksamkeitserwartung der männlichen Schüler auf. Diese Differenz ist allerdings nicht signifikant ($t(43)=1,28$; $p=0,207$). Eine Varianzanalyse mit Messwiederholung zeigt signifikante Unterschiede in der zeitlichen Entwicklung der Selbstwirksamkeitserwartung zwischen männlichen und weiblichen Schülern (Interaktion Zeit x Geschlecht: $F(2,86)=3,27$; $p<0,05$; $\eta^2=0,071$). Während sich die planungsbezogene Selbstwirksamkeit der weiblichen Schüler während des Laborbesuches nicht verändert, steigert sich die der männlichen Schüler durch den Laborbesuch deutlich und sinkt bis zum Follow-up-Test wieder fast auf Pretestniveau ab (Tab. 137 & Abb. 86). Ein signifikanter Unterschied zwischen männlichen Schülern der Versuchs- und männlichen Schülern der Kontrollgruppe in Hinblick auf die Entwicklung der Selbstwirksamkeitserwartung kann allerdings nicht belegt werden ($F(2,74)=0,979$; $p=0,380$).

Tabelle 137: Ergebnisse der Nebenstudie -Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der weiblichen und männlichen Schüler der Versuchsgruppe in der FSWEx.V2-Subskala Planung im Pre-, Post- und Follow-up-Test

Messzeitpunkt	Geschlecht	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	weiblich	1,62	0,52	30
	männlich	1,92	1,06	15
Posttest	weiblich	1,56	0,52	30
	männlich	2,24	1,11	15
Follow-up-Test	weiblich	1,54	0,58	30
	männlich	2,00	0,98	15

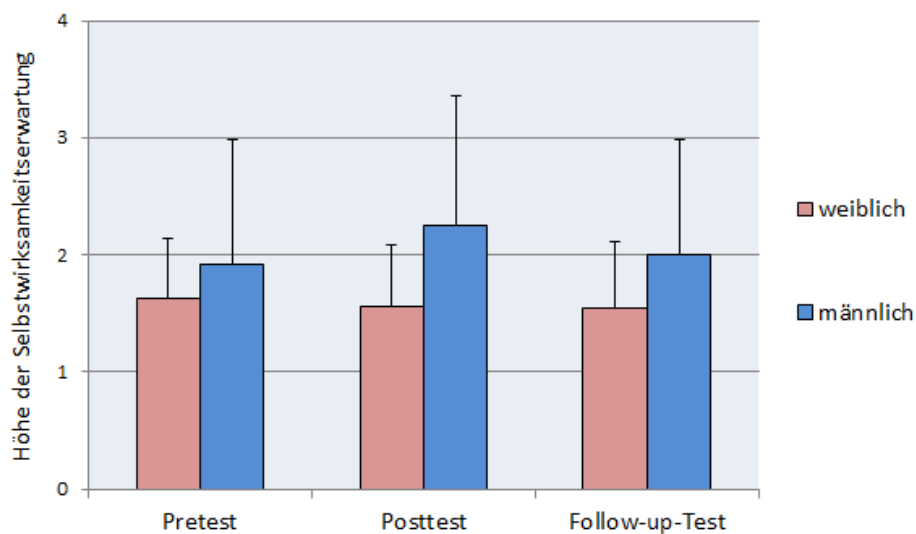


Abbildung 86: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Planung der männlichen und weiblichen Schüler der Versuchsgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

3.2.3.2 FSWEx.V2 – Subskala Durchführung

Auch in der Subskala Durchführung liegen zum Zeitpunkt des Pretests keine signifikanten Unterschiede zwischen Versuchs- und Kontrollgruppe vor ($t(130)=0,95$; $p=0,344$). Ihre Durchführungskompetenz schätzen Schüler beider Gruppen relativ hoch ein. Eine bedeutsame zeitliche Veränderung der Selbstwirksamkeitserwartung ist in der Kontrollgruppe nicht gegeben (Haupteffekt Zeit: $F(2,172)=2,161$; $p=0,118$). Im Gegensatz zur Subskala Planung unterscheiden sich Versuchs- und Kontrollgruppe in der zeitlichen Entwicklung der durchführungsbezogenen Selbstwirksamkeitserwartung über die drei Messzeitpunkte aber deutlich (Interaktion Zeit x Treatment: $F(2,260)=7,46$; $p\leq 0,001$; $\eta^2=0,054$). Vom Pre- zum Posttest ist gegenüber der Kontrollgruppe eine höchst signifikante Steigerung der Selbstwirksamkeitserwartung festzustellen ($F(1,130)=15,71$; $p\leq 0,001$; $\eta^2=0,108$).

Tabelle 138: Ergebnisse der Nebenstudie -Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der Versuchs- und Kontrollgruppe in der FSWEx.V2-Subskala Durchführung im Pre-, Post- und Follow-up-Test

Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Versuchsgruppe	2,75	0,59	45
	Kontrollgruppe	2,65	0,55	87
Posttest	Versuchsgruppe	3,01	0,64	45
	Kontrollgruppe	2,56	0,67	87
Follow-up-Test	Versuchsgruppe	2,80	0,63	45
	Kontrollgruppe	2,57	0,71	87

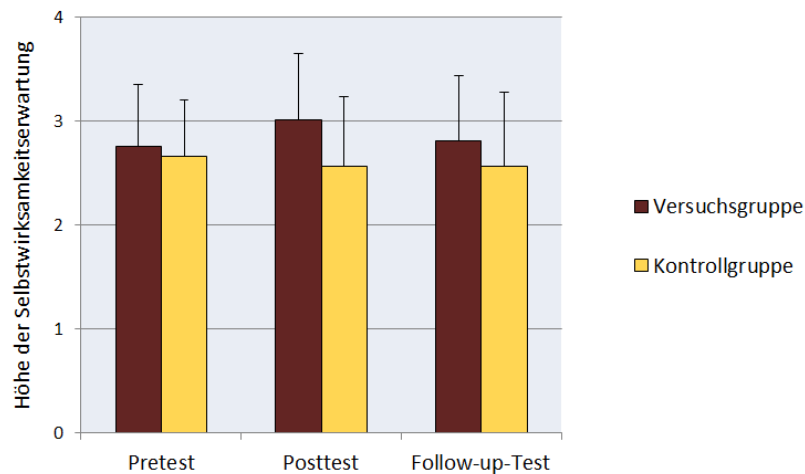


Abbildung 87: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Durchführung der Versuchs- und Kontrollgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

Diese geht allerdings bis zum Follow-up-Test wieder deutlich zurück, so dass vom Pre- zum Follow-up-Test nur eine tendenzielle nachhaltige Steigerung zu belegen ist ($F(1,130)=2,81$; $p=0,096$). Der Besuch des BeLL Bio-Kurses Genetischer Fingerabdruck führt somit kurzfristig zu einer Steigerung der experimentbezogenen Selbstwirksamkeitserwartung in der Subskala Durchführung (Tab. 138 & Abb. 87).

Derartige Unterschiede zwischen den Geschlechtern, wie dies bei der planungsbezogenen Selbstwirksamkeitserwartung der Fall war, finden sich bei der Subskala Durchführung nicht. Weder im Pretest ($t(43)=0,021$; $p=0,779$) noch in der Entwicklung der Selbstwirksamkeitserwartung ($F(2,86)=0,087$; $p=0,917$) unterscheiden sich weibliche und männliche Schüler. Beide Geschlechter steigern also durch den Laborbesuch gleichermaßen ihre durchführungsbezogene Selbstwirksamkeitserwartung.

3.2.3.3 FSWEx.V2 – Subskala Auswertung

Auch für die FSWEx.V2-Subskala Auswertung sind die geringen im Pretest zu beobachtenden Unterschiede zwischen Versuchs- und Kontrollgruppe statistisch nicht bedeutsam ($t(129)=1,87$; $p=0,064$). Eine zeitliche Veränderung der Kontrollgruppe ist nicht gegeben (Haupteffekt Zeit: $F(2,170)=0,31$; $p=0,736$), während sich die auswertungsbezogene Selbstwirksamkeitserwartung der Schüler in der Vergleichsgruppe im Posttest augenscheinlich nach dem Laborbesuch kurzfristig erhöht (Tab. 139 & Abb. 88). Ein signifikanter Unterschied zwischen beiden Gruppen kann im Rahmen einer Varianzanalyse mit Messwiederholung über die drei Messzeitpunkte nicht festgestellt werden (Interaktion Zeit x Treatment: $F(2,258)=1,75$; $p=0,176$).

Tabelle 139: Ergebnisse der Nebenstudie -Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der Versuchs- und Kontrollgruppe in der FSWEx.V2-Subskala Auswertung im Pre-, Post- und Follow-up-Test

Messzeitpunkt	Treatmentgruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	Versuchsgruppe	2,43	0,61	45
	Kontrollgruppe	2,22	0,59	86
Posttest	Versuchsgruppe	2,54	0,76	45
	Kontrollgruppe	2,21	0,62	86
Follow-up-Test	Versuchsgruppe	2,42	0,60	45
	Kontrollgruppe	2,25	0,68	86

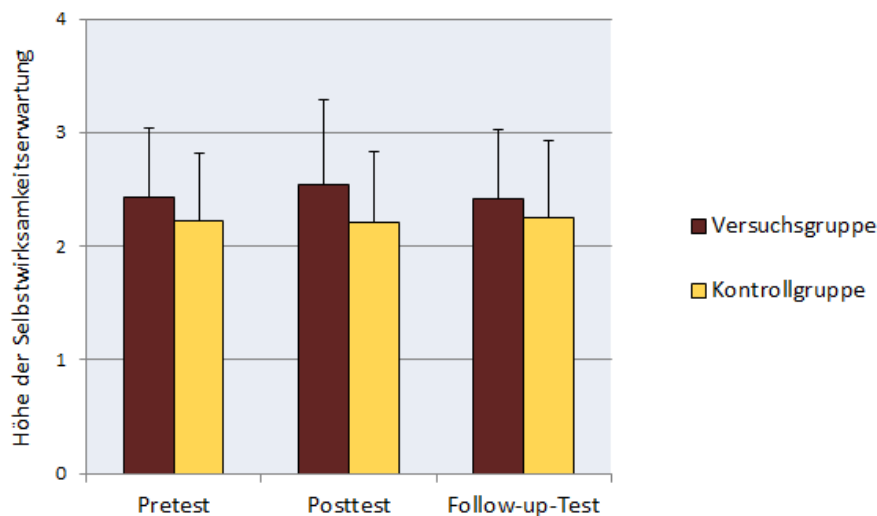


Abbildung 88: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Auswertung der Versuchs- und Kontrollgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

In der Subskala Auswertung liegt im Pretest kein nennenswerter Unterschied zwischen weiblichen und männlichen Schülern vor ($t(43)=0,349$; $p=0,729$). Ähnlich wie in der Subskala Planung ist auch hier zu beobachten, dass sich die auswertungsbezogene Selbstwirksamkeitserwartung der weiblichen Schüler nicht verändert, während die der männlichen Schüler augenscheinlich kurzfristig nach dem Laborbesuch zunimmt (Tabelle 140 & Abbildung 89). Ein signifikanter Unterschied in der zeitlichen Veränderung ist zwischen den Geschlechtern hier allerdings nicht zu belegen (Interaktion Zeit x Geschlecht: $F(2,86)=1,85$; $p=0,164$).

Tabelle 140: Ergebnisse der Nebenstudie -Mittelwerte (*M*) und Standardabweichungen (*SD*) sowie Stichprobengrößen (*N*) der weiblichen und männlichen Schüler der Versuchsgruppe in der FSWEx.V2-Subskala Auswertung im Pre-, Post- und Follow-up-Test

Messzeitpunkt	Geschlecht	<i>M</i>	<i>SD</i>	<i>N</i>
Pretest	weiblich	2,40	0,59	29
	männlich	2,47	0,66	16
Posttest	weiblich	2,43	0,72	29
	männlich	2,74	0,80	16
Follow-up-Test	weiblich	2,40	0,59	29
	männlich	2,46	0,62	16

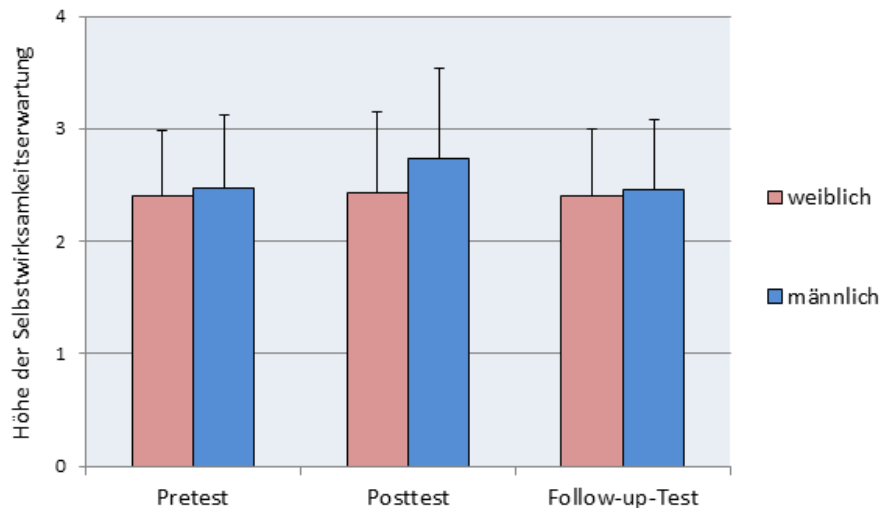


Abbildung 89: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Auswertung der männlichen und weiblichen Schüler der Versuchsgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Skalenmaximum=4)

Bisher wurden ausschließlich die Resultate der Schülerbefragungen in Haupt und Nebenstudie berichtet. Im folgenden Abschnitt werden die Ergebnisse der Lehrerbefragung präsentiert.

3.2.4 Lehrerbefragung

Schülerlabore betreiben einen enormen finanziellen und organisatorischen Aufwand, um Schülern am außerschulischen Lernort Labor die Möglichkeit zum Experimentieren zu geben. Die Notwendigkeit von Schülerlaboren betont die nachfolgende Feststellung. Auf die Frage, welche Probleme Lehrer sehen, wenn sie in der Schule experimentieren möchten, antworteten diese wie in Abbildung 90 dargestellt:

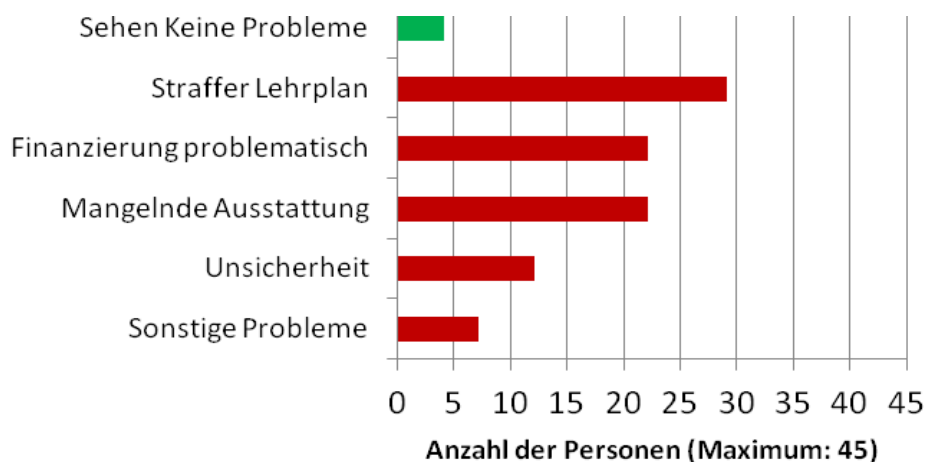


Abbildung 90: Äußerungen von Lehrern zu Problemen beim Experimentieren am Lernort Schule. Angegeben sind absolute Häufigkeiten

Fast zwei Drittel der 45 befragten Lehrer kreuzten im Fragebogen einen zu straffen Lehrplan als Grund für die geringe Experimentierhäufigkeit im Unterricht an. Materielle Gründe, wie einen zu geringen Schuletat sowie die mangelnde Ausstattung des Unterrichtsraumes wurden von etwa der Hälfte der befragten Personen genannt. Gut ein Viertel der Lehrer gaben ihre eigene Unsicherheit als Grund dafür an, einige Experimente nur ungerne durchzuführen. Sonstige Probleme – häufig ohne nähere Beschreibung – wurden nur von einem geringen Anteil geäußert. Mit vier Personen ist der Anteil an Lehrern, der keine Probleme für das Experimentieren in der Schule erkennt, auffällig klein.

Dass der relativ hohe Anteil unsicherer Lehrer sich unmittelbar in ihrer subjektiv empfundenen Selbstsicherheit sowie in den Subskalen des FSWEx.V1 widerspiegelt, zeigen die beiden folgenden Abbildungen. Hier wurden die Lehrpersonen, die explizit die eigene Unsicherheit als problematisch erkannten den übrigen „selbstsicheren“ Lehrern gegenübergestellt. Wie erwartet, zeichnen sich die selbstsicheren Lehrer durch eine höchst signifikant höhere, subjektiv empfundene Selbstsicherheit im Experimentieren gegenüber den unsicheren Lehrern aus ($t(43)=7,41$; $p \leq 0,001$, $\omega^2=0,545$; vgl. Tabelle 141 & Abb 91).

Tabelle 141: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) der selbstsicheren und unsicheren Lehrpersonen in der subjektiv empfundenen Sicherheit

Subjektiv empfundene Selbstsicherheit	<i>M</i>	<i>SD</i>	<i>N</i>
Selbstsichere Lehrer	3,42	0,64	33
Unsichere Lehrer	1,75	0,75	12

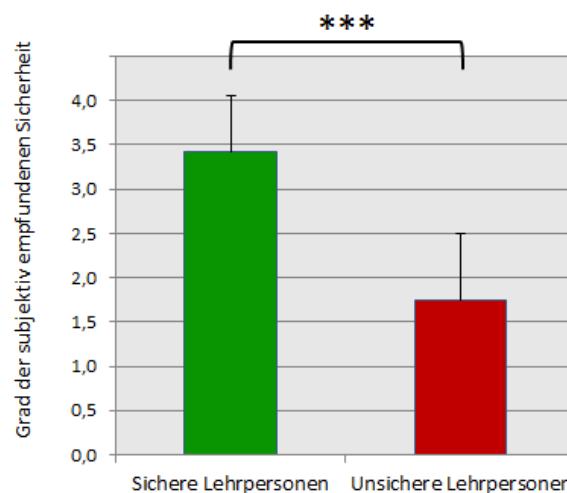


Abbildung 91: Mittelwerte und Standardabweichungen der selbstsicheren und unsicheren Lehrpersonen in der subjektiv empfundenen Selbstsicherheit (Skalenmaximum=4, *= $p \leq 0,001$)**

Ein differenzierteres Bild geben die Messwerte des FSWEx.V1. Wie Abbildung 92 zeigt, schätzen die selbstsicheren Lehrer sowohl ihre Planungs- ($t(43)=5,11$; $p \leq 0,001$, $\omega^2=0,358$), Durchführungs- ($t(43)=4,85$; $p \leq 0,001$, $\omega^2=0,334$) als auch ihre Auswertungskompetenz höchst signifikant ($t(43)=4,21$;

$p \leq 0,001$, $\omega^2 = 0,271$) höher ein als die Vergleichsgruppe (Tabelle 142 & Abbildung 92). Entsprechendes spiegeln die höchst signifikanten Korrelationen zwischen den FSWEx.V1-Subskalen und der Messdimension subjektiv empfundene Selbstsicherheit wieder. Diese liegen mit $r_{\text{Planung/Selbstsicherheit}(45)} = 0,64$; $p \leq 0,001$, $r_{\text{Durchführung/Selbstsicherheit}(45)} = 0,83$; $p \leq 0,001$ und $r_{\text{Auswertung/Selbstsicherheit}(45)} = 0,65$; $p \leq 0,001$ im mittleren bis hohen Bereich.

Tabelle 142: Mittelwerte (M), Standardabweichungen (SD) und Stichprobengrößen (N) der FSWEx.V1-Subskalen für die selbstsicheren und unsicheren Lehrpersonen.

Subskala	Gruppe	M	SD	N
Planung	Selbstsichere Lehrer	2,84	0,57	33
	Unsichere Lehrer	1,81	0,69	12
Durchführung	Selbstsichere Lehrer	3,52	0,38	33
	Unsichere Lehrer	2,20	0,92	12
Auswertung	Selbstsichere Lehrer	3,19	0,67	33
	Unsichere Lehrer	2,20	0,69	12

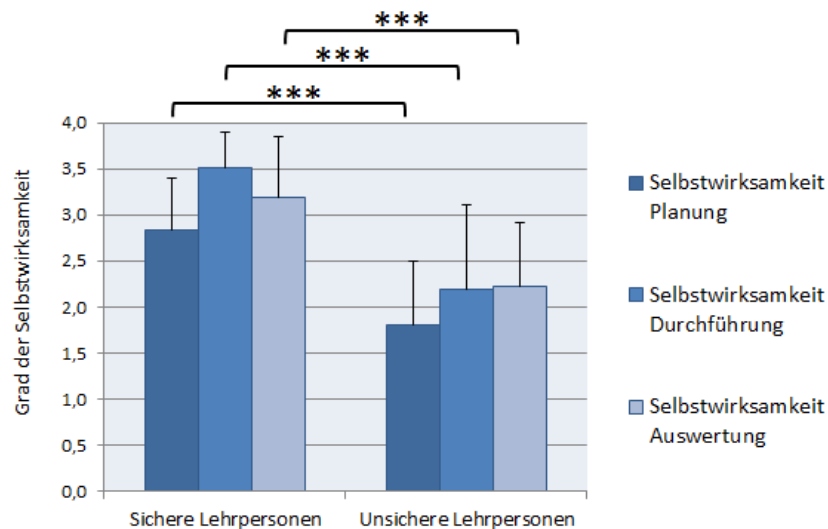


Abbildung 92: Grafische Darstellung der Mittelwerte und Standardabweichungen der selbstsicheren und unsicheren Lehrpersonen in den Selbstwirksamkeits-Subskalen des FSWEx.V1 (Skalenmaximum=4, *= $p \leq 0,001$).**

Während die unsicheren Lehrer – wie zuvor dargestellt – ihre eigenen Fähigkeiten im Experimentieren als relativ gering einstufen, erkennen diese eine ähnlich hohe persönliche Relevanz von Schülerlaborkursen wie die sicheren Lehrpersonen ($t(43) = 7,41$; $p > 0,05$).

Tabelle 143: Mittelwerte (M), Standardabweichungen (SD) und Stichprobengrößen (N) der subjektiv empfundenen Selbstsicherheit von selbstsicheren und unsicheren Lehrpersonen.

Wertebezogenes Interesse	M	SD	N
Selbstsichere Lehrer	3,58	0,51	32
Unsichere Lehrer	3,44	0,89	12

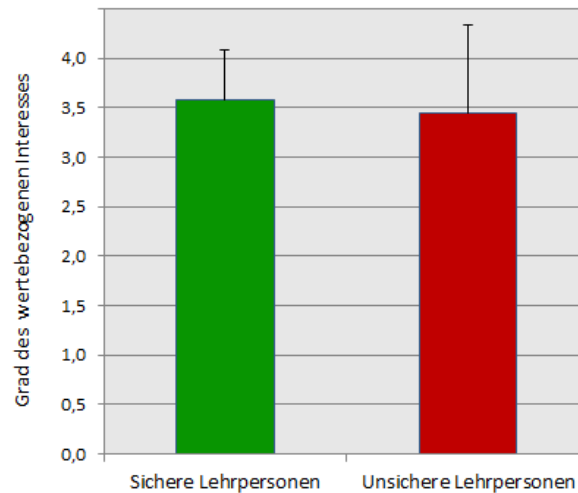


Abbildung 93: Grafische Darstellung der Mittelwerte und Standardabweichungen der selbstsicheren und unsicheren Lehrpersonen im wertbezogenen Interesse an Schülerlaborkursen (Skalenmaximum=4).

Ob die subjektiv empfundene Experimentierkompetenz auch im Zusammenhang steht mit der Häufigkeit, in der die Lehrpersonen Experimente in ihrem Unterricht einsetzen, soll im Folgenden geprüft werden. Augenscheinlich deutet sich an, dass die besonders häufig experimentierenden Lehrpersonen ihre Experimentierkompetenz in den Subskalen des FSWE.V1 höher einschätzen. Lehrer, die dagegen seltener als einmal im Schuljahr experimentieren, schätzen ihre Kompetenz in allen Subskalen sehr gering ein. Deutliche Unterschiede zwischen den einmal monatlich und einmal im Schuljahr experimentierenden Personen sind nicht ersichtlich (Tab. 144, Abb. 94).

Tabelle 144: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) der FSWE.V1-Subskalen in Abhängigkeit von der Experimentierhäufigkeit

Subskala	Experimentierhäufigkeit	<i>M</i>	<i>SD</i>	<i>N</i>
Planung	ca. 1x pro Woche	3,67	0,00	2
	ca. 1x pro Monat	2,50	0,64	23
	ca. 1x pro Schuljahr	2,64	0,69	14
	Seltener	0,83	0,24	2
Durchführung	ca. 1x pro Woche	3,67	0,00	2
	ca. 1x pro Monat	3,29	0,54	23
	ca. 1x pro Schuljahr	3,20	0,82	14
	Seltener	0,67	0,47	2
Auswertung	ca. 1x pro Woche	4,00	0,00	2
	ca. 1x pro Monat	3,00	0,64	23
	ca. 1x pro Schuljahr	2,88	0,94	14
	Seltener	1,50	0,24	2

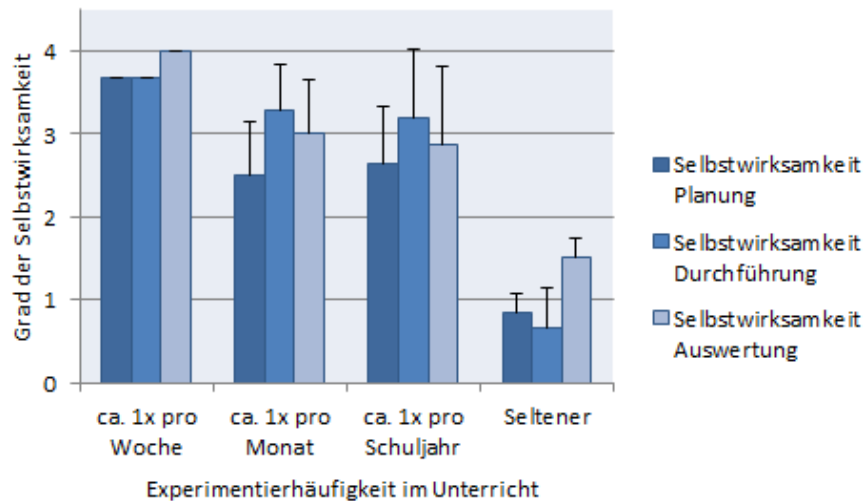


Abbildung 94: Grafische Darstellung der Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 in Abhängigkeit von der Experimentierhäufigkeit der Lehrpersonen im Unterricht (Skalenmaximum=4)

Die geringen Korrelationen zwischen der Experimentierhäufigkeit und den Subskalen Planungskompetenz ($r_{\text{Kendall-Tau}}=-0,183$; $p=0,082$) und Durchführungskompetenz ($r_{\text{Kendall-Tau}}=-0,171$; $p=0,103$) bestätigen diesen Zusammenhang statistisch nicht. Die Korrelation der Subskala Auswertungskompetenz ($r_{\text{Kendall-Tau}}=-0,247^8$; $p=0,031$) ist auf dem Niveau von $p \leq 0,05$ signifikant. Zu beachten ist, dass die Gruppengrößen der sehr seltenen und sehr häufigen Experimentierer nur mit zwei Personen besetzt sind und daher hier keine gesicherten Aussagen zu treffen sind.

Tabelle 145 und Abbildung 95 sollen zeigen, ob die empfundene Experimentierfähigkeit von Lehrkräften im Zusammenhang mit der von ihnen im Schulunterricht präferierten Experimentierform steht. Augenscheinlich stufen die Lehrpersonen, die Schülerexperimente bevorzugen, ihre Kompetenz Experimente zu planen ($F(1,41)=3,26$; $p=0,078$), durchzuführen ($F(1,41)=3,88$; $p=0,056$) und auszuwerten ($F(1,41)=2,80$; $p=0,102$) tendenziell, wenn auch nicht signifikant, höher ein als Lehrpersonen, die Lehrerexperimente bevorzugen.

Tabelle 145: Mittelwerte (*M*), Standardabweichungen (*SD*) und Stichprobengrößen (*N*) der FSWEx.V1-Subskalen in Abhängigkeit von der Präferenz, im Unterricht eher Lehrer- bzw. Schülerversuche durchzuführen

Subskala	Gruppe	<i>M</i>	<i>SD</i>	<i>N</i>
Planung	Lehrerversuche	2,11	1,026	6
	Schülerversuche	2,68	0,655	37
Durchführung	Lehrerversuche	2,72	1,063	6
	Schülerversuche	3,32	0,613	37
Auswertung	Lehrerversuche	2,50	0,937	6
	Schülerversuche	3,05	0,722	37

⁸ Die Korrelation ist negativ, da mit zunehmender Seltenheit des unterrichtlichen Experimentierens höhere Zahlen in der Variablenkodierung vergeben wurden.

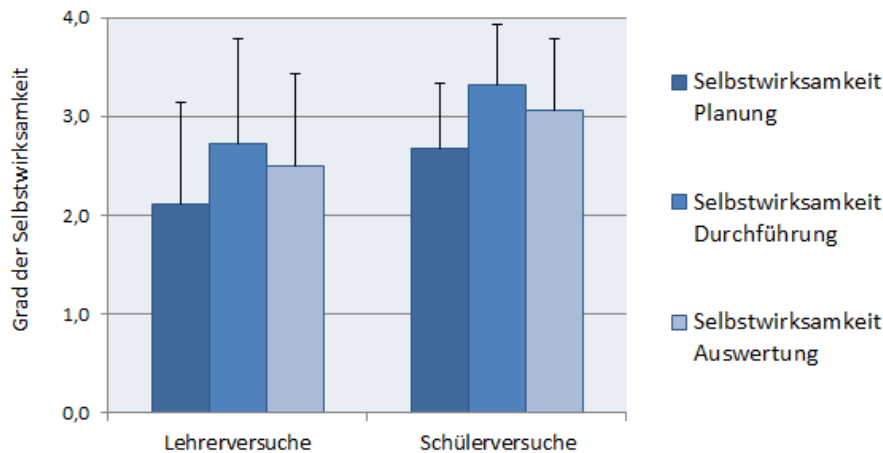


Abbildung 95: Grafische Darstellung der Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 für Lehrpersonen, die häufiger Lehrerversuche bzw. Schülerversuche durchführen (lassen), (Skalenmaximum=4)

Interessant ist in diesem Zusammenhang auch die Betrachtung der Anteile selbstsicherer und unsicherer Lehrpersonen in der Gruppe der Lehrerversuche- bzw. Schülerversuche-Präferierenden (Tab. 146). Zunächst wird ersichtlich, dass der überwiegende Anteil der Lehrpersonen angibt, im schulischen Unterricht bevorzugt Schülerversuche durchzuführen. Mit 27,3 % findet sich allerdings knapp ein Drittel der unsicheren Lehrpersonen in der Gruppe wieder, die Lehrerversuche bevorzugen. Dagegen befinden sich in dieser Gruppe nur 9,4% der selbstsicheren Lehrer. Der exakte Test nach Fisher⁹ belegt allerdings keinen signifikanten Zusammenhang zwischen Experimentierpräferenz und der eigenen Einstufung der Selbstsicherheit ($p=0,164$).

Tabelle 146: Die Kreuztabelle zeigt die Experimentierpräferenz selbstsicherer und unsicherer Lehrpersonen

Gruppe	Experimentierpräferenz		Gesamt
	Lehrerversuche	Schülerversuche	
Selbstsicher	N=3 (9,4%)	N=29 (90,6%)	N=32 (100%)
Unsicher	N=3 (27,3%)	N=8 (72,7%)	N=11 (100%)

Nachdem bisher die zentralen Ergebnisse der vorliegenden Arbeit berichtet wurden, werden diese im folgenden Abschnitt diskutiert.

⁹ Der exakte Test nach Fisher wurde gewählt, da ein χ^2 -Test aufgrund der in zwei Zellen der Kreuztabelle auftretenden Fallzahl <5 nicht durchgeführt werden konnte (vgl. Bühl 2010).

4 Diskussion

Ein Bestandteil der vorliegenden Arbeit war es, die beiden lehrplanrelevanten Themenbereiche Molekularbiologie und Zellbiologie in experimentellen Laborkursen umzusetzen. Daher wurden zunächst zwei komplette Kurstage für das Schülerlabor Bell Bio zu den Themen „Genetischer Fingerabdruck – Dem Täter auf der Spur“ (Molekularbiologie) und „Außer Atem – Sportliche Enzyme“ (Zellbiologie/Zellphysiologie) konzipiert. Der letztgenannte Kurs wurde im Rahmen einer Bachelor-Thesis als Vorstudie bereits einmal durchgeführt (Aepfelbach 2011). Der im fachdidaktischen Teil ebenfalls analysierte Laborkurs „Kleine Einzeller ganz groß“ (Zellbiologie) wurde im Rahmen einer Bachelor-Thesis konzipiert (Kaminska 2010).

Das Konzept des Bell Bio sieht vor, aktuelle und forschungsrelevante Kurse zu veranstalten und Schülern die Auseinandersetzung mit modernen Labor-Methoden und Arbeitsabläufen zu bieten. Die fachwissenschaftlichen Optimierungen für die oben genannten Kurse werden nachfolgend diskutiert. Die Diskussion der Ergebnisse der fachdidaktischen Evaluierung dieser Kurse findet sich im Anschluss.

4.1 Fachwissenschaftliche Untersuchungen

4.1.1 Der Genetische Fingerabdruck

Da Fernsehserien wie CSI eine Hochzeit an Einschaltquoten erleben und diese insbesondere von Jugendlichen gesehen werden, scheint ein großes Interesse an Themen der Rechtsmedizin in dieser Altersgruppe zu bestehen (Keuneke et al. 2010). Der Genetische Fingerabdruck findet vielfältige Verwendung, z.B. in der Verwandtschaftsforschung (Kayser & Sajantila 2001; Naito et al. 2003), dem DNA-Barcoding der Biodiversität (Goldstein & De Salle 2010) oder der Forensischen Kriminalistik (Ballantyne et al. 2010; Morling 2009). Die Erstellung eines Genetischen Fingerabdruckes erfordert die Nutzung vielfältiger, aktueller molekularbiologischer Methoden, Geräte und Techniken. Im Rahmen dieses Laborkurses erhalten die Schüler die Chance, ihre Medien-bedingten Fehlkonzepte in realistischere Vorstellungen zu überführen. Dies gelingt im Kontext des Schülerlabors umso eher, da entsprechende Methoden im Biologieunterricht, sei es bedingt durch fehlende Gerätschaften oder aus Mangel an praktischer Experimentiererfahrung der Lehrkraft, kaum zu realisieren sind. Im Schülerlabor BeLL Bio sind sowohl die Geräte als auch die laborpraktischen Kompetenzen vorhanden. Die hier vorgestellten fachlichen Arbeiten tragen zudem dazu bei, auch die Methoden und Arbeitsabläufe für den Einsatz im Schülerlabor zu optimieren.

DNA-Isolierung: Bei der Übertragung aktueller Forschung auf das Niveau von Schülern müssen für die Praxis solche Methoden etabliert werden, die einen nahezu reibungslosen und vor allem unbedenklichen Ablauf der Experimente garantieren. Extraktionen mit z.B. Phenol-Chloroform scheiden somit aus (vgl. DGUV 2010). Der DNA kommt dabei als zentrales Makromolekül der Molekularen Genetik eine besondere Bedeutung zu. Die Isolierung der DNA ist zunächst einmal eine wichtige Voraussetzung, um überhaupt molekulargenetisch arbeiten zu können. Gleichzeitig kann man anhand der einzelnen Schritte bei der Extraktion sehr gut auf die biochemischen Eigenschaften der DNA eingehen, da die Kenntnis dieser Zusammenhänge eine Voraussetzung für die korrekte Handhabung der Proben ist. Im Rahmen der DNA-Isolierung können also theoretisches Vorwissen und praktische Erfahrung der Schüler sehr gut kombiniert werden. Es sollte also eine Methode gefunden werden, bei der Schritte wie Solubilisierung, Trennung oder Fällung einzelne Handlungen oder Chemikalien bedingten und nicht – wie in einer Black Box – in undefinierten Lösungen und Methoden vorliegen. Drei verschiedene Methoden wurden daher experimentell angewandt und für einen Einsatz im Schülerlabor getestet. Das im Labor etablierte DNeasy Plant Mini Kit führt zu sehr guten Ausbeuten an sauber extrahierter DNA (Abb. 37, Punkt 3.1.1.1). Leider erfüllt es mehrere sehr wichtige Kriterien nicht: Die Schritte des Kits sind für Schüler nicht immer nachvollziehbar, da im Unterricht häufig traditioneller Methoden der DNA-Isolierung behandelt werden. Zudem ist die genaue Zusammensetzung mancher Lösungen nicht bekannt, was die Nachvollziehbarkeit einzelner Schritte erschwert. Weiterhin wäre es wünschenswert, die Probe in flüssigem Stickstoff zu mörsern, um einen idealen Zellaufschluss zu erreichen. Dieses könnte eine potentielle Gefahrenquelle für Schüler darstellen. Das Kit funktioniert auch mit einem stärkeren Lysepuffer, bringt dann aber keine so guten Ausbeuten an reiner DNA. Ein deutliches Ausschlusskriterium für die Verwendung im BeLL Bio stellten außerdem die relativ hohen Kosten dar.

Eine sowohl relativ ungefährliche, als auch kostengünstige Variante ist der Aufschluss der Buccalzellen mittels Lysepuffer und der folgenden Kaliumacetat-Isopropanol-Fällung der Proteine und DNA. Dieses Verfahren ist geeignet, im Laborkurs von den Schülern selbst vorgenommen zu werden, da die einzelnen Schritte gut nachvollziehbar sind. Die durch die Inkubationszeiten und Zentrifugierschritte entstehenden Pausen können von den Schülern genutzt werden, um weitere Arbeitsschritte selbständiger zu planen und gleichzeitig erleben zu können, wie der Laboralltag mit ineinander verschachtelten Arbeitsabläufen typischerweise aussieht. Die DNA-Ausbeute dieser Methode ist allerdings geringer, so dass diese Methode sich als ungeeignet erwies, um im Rahmen der Laborkursvorbereitung DNA-Proben der „Tatverdächtigen“ zu isolieren.

Die Methode, die für die Vorbereitung der im BeLL Bio eingesetzten DNA-Proben etabliert wurde, ist die Extraktion der DNA mit Hilfe von Chelex. Diese ist zwar etwas zeitaufwändiger, als die Kit-

Isolierung, dafür ist sie aber kostengünstiger und bietet eine höhere Ausbeute, als die Kaliumacetat-Isopropanol-Extraktion (Abb. 37, Spuren 3+4). Das Ionenaustauscherharz Chelex ist ein Styrol-Divinylbenzol-Copolymer mit Iminodiacetat-Gruppen. Diese sind in der Lage, bivalente Ionen wie Mg^{2+} oder Ca^{2+} zu komplexieren (http://www.bio-rad.com/webmaster/pdfs/9184_Chelex.pdf, geprüft am 10.01.12, Goodwin et al. 2011). Damit können viele Enzyme, die auf bivalente Ionen angewiesen sind, wie z.B. DNasen, ihre katalytische Aufgabe nicht wahrnehmen und sind effektiv blockiert. Zudem werden Ca^{2+} -abhängige Strukturkomponenten der Zelle, wie extrazelluläre Strukturproteine und Bestandteile des Cytoskeletts durch das Abfangen von Ca^{2+} geschwächt, was zu einer Destabilisierung der Zellen führt. Die im nächsten Schritt zugesetzte Proteinase K ist Ca^{2+} -unabhängig und führt über den Abbau dieser Proteine zur Lyse der Zelle, so dass die DNA freigesetzt werden kann. Damit die Proteinase K in der sich anschließenden PCR nicht die eingesetzten Enzyme verdaut, muss sie durch Hitze-Denaturierung inaktiviert werden (<http://sundoc.bibliothek.uni-halle.de/diss-online/07/07H096/t4.pdf>, geprüft am 10.01.12). Da die *Taq*-Polymerase in der PCR Mg^{2+} -abhängig arbeitet, müssen auch die Chelexkügelchen vor der PCR entfernt werden. Ein Problem der Chelex-Extraktion ist, dass die isolierte DNA häufig nur noch niedermolekular vorliegt, d.h. es liegen nur relativ kurze DNA-Stränge für die PCR vor. Da im vorliegenden Fall aber nur niedermolekulare VNTRs untersucht wurden, ist das kein störender Faktor (Walsh et al. 1991; Wiegand et al. 1993; Singer-Sam et al. 1989, Jung 2002, Mülhardt 2009).

Anfärbung der DNA-Banden im Agarosegel: Im Anschluss an die optimierte DNA-Extraktion mussten die VNTRs in der PCR amplifiziert, in der Gelelektrophorese nach der Größe aufgetrennt und mit einem geeigneten, das heißt effektiven, kostengünstigen und ungefährlichen Verfahren angefärbt werden. Für den Einsatz im BeLL Bio erwies sich der Fluoreszenzfarbstoff Stain G als geeignet, wie anhand des Agarosegels überprüft werden kann (Abb. 38, B). Stain G emittiert eine grüne Fluoreszenz mit einem Maximum bei ca. 530 nm, wenn es an DNA gebunden vorliegt. Es besitzt eine ähnliche Sensitivität wie Ethidiumbromid und kann direkt bei der Zubereitung zur Agarose-Gellösung hinzugegeben werden (http://www.serva.de/www_root/documents/39803_d.pdf, geprüft am 10.01.12).

Das in die DNA interkalierende Ethidiumbromid (Abb. 38, A) ist ein stark giftiger Phenanthridin-Farbstoff, der Nukleinsäuren anfärbt. Mit diesem dürfen Schüler in der Schule nicht selbst arbeiten (DGUV 2010). Entsprechend wurde auf eine Verwendung im BeLL Bio verzichtet. Der Farbstoff GelRedTM ist grundsätzlich auch geeignet (Abb. 38; C) und sehr sensitiv. Er wurde aber aufgrund seines höheren Preises für den dauerhaften Einsatz im BeLL Bio verworfen.

Die drei zuvor genannten Farbstoffe werden erst durch Anregung mit UV-Licht der Wellenlänge 312 nm sichtbar. Als ungeeignet erwies sich der nicht fluoreszierende Farbstoff Methylenblau (Abb. 38;

D). Zwar ist er relativ kostengünstig und relativ ungefährlich, allerdings benötigt er zu große Mengen an DNA, um diese zuverlässig detektieren zu können. Gerade bei genetischen Arbeiten im Schülerlabor kann es durch unsaubere Extraktionen oder leicht fehlerhaft angesetzte PCR-Reaktionen dazu kommen, dass nur kleine Mengen DNA in den Banden vorhanden sind. Ein empfindlicher, verhältnismäßig unbedenklicher und relativ kostengünstiger Farbstoff, wie er mit Stain G vorliegt, ist daher von Vorteil.

PCR-Protokolle: Nachdem nun die Extraktionsmethoden und die Farbstoffe ausgewählt waren, mussten die Bedingungen für die PCR-Reaktionen ermittelt werden. Eine Gradienten-PCR ist eine ideale Möglichkeit, um verschiedene Temperaturen in einem PCR-Lauf zu testen (Zhang & Xing 2009). Entscheidend für eine erfolgreiche PCR ist neben der Aufspaltung der doppelsträngigen DNA (dsDNA) in Einzelstrang-DNA (ssDNA) durch Hitzedenaturierung die Anlagerung der Primer an die ssDNA. Dieser als Annealing bezeichnete Prozess benötigt in Abhängigkeit von der Basenzusammensetzung und der Länge eine geeignete Temperatur (Hwang et al. 2003). Diese liegt meistens etwas tiefer als die zu berechnende Schmelztemperatur der Primer-DNA. Es ist allerdings bekannt, dass diese ideale Temperatur nicht immer zu den besten Ergebnissen führt, da Wechselwirkungen zwischen den Primern und der eingesetzten DNA vorkommen (Hwang et al. 2003). So können auch zu kurze oder zu lange Reaktionsprodukte entstehen, die unspezifisch sind und nicht das gewünschte Produkt darstellen. Außerdem sind Reaktionen zwischen den Primern selbst oder Haarnadelstrukturen bekannt (Abd-Elsalam 2003), die das Annealing stören. Daher wurde für die Primerpaarungen ein Temperaturgradient eingesetzt, der in kleinen Schritten für Reaktionsgefäße festgelegter Position, die Temperatur variiert (Zhang & Xing 2009). Auf diese Weise wurden die Annealingtemperaturen experimentell bestimmt und nicht nur berechnet. Die Banden im Agarosegel zeigen insbesondere für den VNTR D1S80 eine offensichtlich gut gewählte Annealing-Temperatur (Abb. 40).

Eine weitere Rolle spielen auch die Zeiten, die für die Phasen einer PCR angesetzt werden. Die primäre Denaturierung muss lang genug andauern, dass sich auch ein DNA-Doppelstrang mit einem hohen Gehalt an Guanin und Cytosin (GC-Gehalt) korrekt in zwei Einzelstränge trennt. Ein hoher GC-Gehalt in der DNA führt aufgrund der stärkeren Bindungen durch Wasserstoffbrücken und einer höheren thermodynamischen Stabilität zu einem temperatur-stabileren DNA-Doppelstrang (Yakovchuk et al. 2006). Bei den hier untersuchten DNA-Abschnitten wurde eine fünfminütige initiale Denaturierung bei 95°C durchgeführt. Die gleichen Bedingungen muss die Denaturierung zu Beginn eines jeden Zyklus erfüllen. Da hier die gebildeten DNA-Stücke noch recht kurz sind, genügt eine Temperatur von 94°C für 1 Minute, um ssDNA zu erhalten. Die Anzahl der Zyklen, bei denen der Denaturierung das Annealing und die Elongation folgen, spielen ebenfalls eine große Rolle. Bei

schwer zu amplifizierenden DNA-Abschnitten kann eine höhere Zyklenzahl zum Erfolg führen. Bei den vorliegenden Untersuchungen musste der VNTR-Locus $I_g - J_H$ mit einer Zyklenzahl von 35 gegenüber sonst 30 Zyklen amplifiziert werden (Decorte et al. 1990). Je nach erwarteter Zusammensetzung und Länge der PCR-Produkte liegen die Zeiten für das Annealing und die Elongation im Sekunden- bis Minutenbereich. Da die hier verwendeten Primer die flankierenden Bereiche der VNTRs abgreifen, müssen sie fest genug binden, dass sie auch bei gegebenenfalls vielen Sequenzwiederholungen das Produkt erbringen, ohne dass die Komplettierung des DNA-Stranges durch die Oligonukleotide frühzeitig abbricht (Horn et al. 1989; Budowle et al. 1991; Ivey et al. 1994; Duncan et al. 1997). Mit der finalen Elongation schließt dann die Anlagerung weiterer Oligonukleotide an die ssDNA ab. Da bei einer Standardeinstellung von 10 min. bei 72°C im Allgemeinen gute Produkte erzielt wurden, wurde diese beibehalten.

Auswahl der VNTRs: In hier vorliegenden Untersuchungen wurde mit DNA-Abschnitten gearbeitet, die nicht Bestandteil der Gene sind. VNTRs sitzen zwar sehr nahe oder auch mitten in Genen (oder auch Introns), sie werden aber nicht transkribiert, sondern dienen wahrscheinlich zur Regulation von Transkription und Expression von Genen (Babushkina & Kucher 2011; Nakamura et al. 1998). VNTRs zeigen einen ausgeprägten Polymorphismus, weswegen sie zur Unterscheidung von Individuen eingesetzt werden können. Ein Großteil der Gesamt-DNA des Menschen besteht aus repetitiven, nicht kodierenden Einheiten (Jasinska & Krzyzosiak 2004). In vorliegenden Untersuchungen folgte eine Analyse von vier hypervariablen DNA-Loci, um ihre Eignung zur Erstellung des Genetischen Fingerabdruckes im Schülerlabor zu testen. Die hier gewählten VNTR-Loci sind solche, die auch in der Rechtsmedizin Verwendung finden. Allerdings sollen hier nicht Wissenschaftler mit den Proben arbeiten, sondern Schüler. Für den Einsatz im Schülerlabor ist daher ein effektives, transparentes und wenig aufwändiges Verfahren notwendig.

Es stellte sich für den hier verfolgten Zweck der VNTR-Locus D1S80 aufgrund seiner Eigenschaften und den hier verwendeten PCR-Bedingungen als ideal heraus. Der auf dem Chromosom 1 liegende Locus ist hypervariabel, kommt also hoch polymorph in Menschen vor (Budowle et al. 1991, Sajantila et al. 1992). Diese Variabilität kommt durch die hohe Anzahl von Allelen zustande. Da bisher 32 verschiedene Allele bekannt sind und zwei verschiedene Allele pro Individuum vorliegen können, ist die Diversität an allelischen Kombinationen enorm hoch (vgl. Das & Mastana 2003). Diese Diversität in Verbindung mit der sehr guten Reproduzierbarkeit der Banden durch die PCR und Gelelektrophorese, macht diesen DNA-Abschnitt sehr geeignet für solche Fragestellungen (Abb. 40 D). Seine gute Variabilität erlaubt eine schnelle und eindeutige Unterscheidung zwischen verschiedenen Individuen. Weiterhin war es in den durchgeführten Untersuchungen bei diesem

VNTR gut möglich zu unterscheiden, ob das Merkmal hetero- oder homozygot vererbt vorliegt (Abb. 40 D).

Auch der Locus D17S30 zeigt eine sehr hohe Diversität (Abb. 40, A) unter den Probanden und auch in der Literatur. Die meisten Menschen besitzen Allele, in abnehmender Häufigkeit dargestellt, mit 4-, 2- oder 3-facher Wiederholung (Buscemi et al. 1994, Ivey et al. 1994). Allerdings ist aufgrund der zahlreichen unspezifischen Banden nach der Gelelektrophorese eine eindeutige Unterscheidung einzelner Individuen hier nicht möglich. Für Experimente im BeLL Bio erscheint dieser Locus unter den erprobten PCR-Bedingungen daher nicht geeignet.

Auch der VNTR-Locus I_G-J_H zeigte keine eindeutig auswertbaren Ergebnisse. Die Wiederholungssequenzen machen hier 50 Basenpaare aus. Sechs bis acht Allele sind bisher bekannt (Decorte et al. 1990, Ruangjirachuporn et al. 2006). Die erwarteten Größen von 470 bis 920 bp (Decorte et al. 1990) konnten in vorliegender Studie nicht reproduziert werden. Eine Zuordnung zur homo- oder heterozygoten Verteilung ist ebenfalls nicht möglich. Auch dieser Locus erscheint unter den erprobten PCR-Bedingungen nicht für den Einsatz im Schülerlaborkurs geeignet.

Der VNTR Locus apoB3 ist auf dem humanen Chromosom 2 lokalisiert, wo er in der Nähe vom 3'-Ende des Gens für das Apolipoprotein B liegt. Seine Core-Sequenz ist zwischen 14 und 16 Basenpaaren lang und liegt meist zwischen 21- und 57-fach wiederholt vor (Frossard et al. 1999, Das & Mastana 2003). In den Experimenten mit diesem VNTR-Locus wurden bei allen Versuchen (auch mit weiteren, hier nicht dokumentierten DNA-Proben) keine größeren Banden detektiert. Die von Decorte et al. (1990) als häufig identifizierten Banden bei 661 bp konnten hier nicht amplifiziert werden. Weiterhin lagen nicht alle Banden in der erwarteten Größenordnung von 300 bis 900bp und es wurden nur PCR-Produkte für wenige Probanden erhalten. Zudem sahen diese bei verschiedenen Probanden sehr ähnlich aus, so dass die Unterscheidung von zwei Personen nicht eindeutig möglich war. Auch dieser VNTR wurde unter den gewählten PCR-Bedingungen daher für die Verwendung im BeLL Bio ausgeschlossen.

Zusammenfassend erwies sich also der VNTR-Locus D1S80 als sehr gut im Schülerlabor einsetzbar, weil er in vielen verschiedenen Allelen auftritt und die gewählten PCR-Bedingungen zu deutlichen Banden im Agarosegel führten. Diese ließen zudem eine relativ eindeutige Unterscheidung zwischen Homo- und Heterozygoten zu.

Vor dem Hintergrund, dass in Schulklassen viele verschiedene Allele von VNTRs zu erwarten sind, sollte ein VNTR-Locus identifiziert werden, der bei allen ethnischen Hauptgruppen Kaukasier, Asiat und Afrikaner nachzuweisen war (die Frequenz des Vorkommens bestimmter Allele kann mit der ethnischen Zugehörigkeit variieren, Deka et al. 1992a/b, Das & Mastana 2003). Diese

Untersuchungen wurden mit nur wenigen freiwilligen, zufällig ausgewählten Probanden durchgeführt und sie erheben in keiner Weise Anspruch darauf, eine statistisch begründete Aussage über Allelfrequenzen in verschiedenen ethnischen Zugehörigkeiten zu treffen. Sie sollen eher eine zufällig mögliche Verteilung von Schülergruppen simulieren, wie sie im Schülerlabor arbeiten könnten. Es war nicht die Absicht der vorliegenden Untersuchungen, Unterschiede in den Ethnien aufzuzeigen. Vielmehr sollte lediglich die Wahrscheinlichkeit des Vorkommens möglichst vieler verschiedener Allele in der kleinen untersuchten Stichprobe erhöht werden. Wie die Gelbilder der PCR-Produkte des D1S80-Abschnittes (Abb. 40 D & 41) zeigten, stellte sich der Nachweis unabhängig von den auftretenden Allelen als unproblematisch dar. Damit sollte also gewährleistet sein, dass auch bei einer hohen Heterogenität der im Schülerlaborkurs vorkommenden Allele, differenzierbare Produkte erhalten werden.

Für die im Schülerlabor durchgeführten DNA-Analysen ist weiterhin zu berücksichtigen, dass DNA-Untersuchungen zu diagnostischen Zwecken nur von einem Arzt vorgenommen werden dürfen (§ 7 Abs. 1 GenDG). Auch im Rahmen der Strafverfolgung sind Untersuchungen, die über die Ermittlung des DNA-Identifizierungsmusters sowie des Geschlechtes hinausgehen, unzulässig (§ 81g Abs.2 StPO). Der im Schülerlaborkurs untersuchte VNTR hat keinerlei diagnostische Funktion und die vorgenommenen Untersuchungen beschränken sich auf die Unterscheidung von zwei Personen. Die DNA, die im Schülerlabor untersucht wird, stammt von freiwilligen Spendern. Die Proben werden anonym behandelt und niemals einer Person zugeordnet. Im Kurs isolierte DNA wird unmittelbar nach Beendigung des Kurses vernichtet. Die Realisierung des Genetischen Fingerabdruckes im Kontext der Überführung eines Tatverdächtigen erscheint vor diesem Hintergrund unproblematisch.

4.1.2 Photometrische Messungen zum Pyruvatabbau durch die Laktatdehydrogenase

Die photometrischen NADH-Messungen wurden in Orientierung an verschiedenen früheren Studien (Galkin et al. 2001) bei $\lambda=366$ nm durchgeführt. Das Absorptionsmaximum von NADH liegt eigentlich bei $\lambda=340$ nm, die Empfindlichkeit der Messung reicht aber für die meisten Anwendungen auch bei $\lambda=366$ nm aus (Rommel et al. 1968). Diese Wellenlänge wird bevorzugt für NADH-Messungen in Photometern mit Wolframlampe verwendet, da die Strahlungsintensität dieses Leuchtmittels bei entsprechender Wellenlänge noch ausreichende Messempfindlichkeit verspricht (Galkin et al. 2001).

Zur Aufnahme der Absorptionsspektren wurden keine Reinstoffe, sondern die in den Erfrischungsgetränken enthaltenen Farbstoffe verwendet, da sie schnelle Verfügbarkeit und zeitökonomische Anwendung in der Vorbereitung der Schülerlaborkurse versprachen. Angaben zu den enthaltenen Farbstoffen konnten der Getränkeverpackung entnommen werden. Das Spektrum

von Powerade Wild Cherry zeigt den für E122 (Azorubin) erwarteten Verlauf mit einer charakteristischen Schulter bei etwa $\lambda=550$ nm (Heidecke et al. 1983). Für die Anthocyane im Getränk Powerade Zero liegt das beobachtete Absorptionsmaximum im sichtbaren Bereich ebenfalls im erwarteten Bereich von $\lambda=465-560$ nm (Watzl 2002). Zwar ist zu erwarten, dass sich das Absorptionsmaximum der Anthocyane unter den pH-Bedingungen in der Küvette verschiebt (vgl. Ebermann & Elmadfa 2011), wie die gelungene photometrische Verfolgung des NADH-Abbaus aber zeigt, scheint dies jedoch unproblematisch.

Für den NADH-Abbau wurde bewusst Laktatdehydrogenase aus der Skelettmuskulatur und nicht aus der Herzmuskulatur verwendet. Im Skelettmuskel wird das M₄-Isoenzym der LDH stärker exprimiert, während im Herz die H₄-Isoform überwiegt (Berg et al. 2007). Das Isoenzym in der Skelettmuskulatur sorgt schon bei niedrigen Pyruvatkonzentrationen für einen schnellen Umsatz zu Laktat, während das Isoenzym in der Herzmuskulatur auf eine schnelle Oxidation von Laktat zu Pyruvat spezialisiert ist (Horn 2009). Damit erscheint das M₄-Isoenzym für den Einsatz im vorliegenden Versuch geeigneter.

Im Rahmen der photometrischen Messungen konnte festgestellt werden, dass weder Powerade Zero noch Powerade Wild Cherry den vollständigen Umsatz von Pyruvat zu Laktat beeinträchtigen. Die beiden Erfrischungsgetränke sind somit geeignet, um den Pyruvatproben eine rote Färbung zu verleihen. Durch mehrfach wiederholte Messungen konnte diese Beobachtung abgesichert werden. Auf eine Bestimmung der Enzymaktivität der LDH wurde verzichtet, da lediglich der vollständige Umsatz des Pyruvats in den Proben gewährleistet werden sollte. Aussagen über eine möglicherweise hemmende Wirkung der Inhaltsstoffe der Erfrischungsgetränke können und sollen daher nicht getroffen werden. Letztlich wurde – unabhängig vom Zusatz der Erfrischungsgetränke – ein vollständiger Pyruvat-Umsatz in etwa fünf Minuten erreicht. Somit scheint das Vorgehen geeignet, um sowohl zeitökonomisch als auch authentisch im Kontext von Schülerlaborkursen umgesetzt zu werden.

4.2 Fachdidaktisch

Die folgende Diskussion bezieht sich zunächst auf die methodische Umsetzung der Evaluierung, um im Anschluss anhand der didaktischen Fragestellungen die inhaltlichen Ergebnisse der Studie zu diskutieren.

4.2.1 Methodendiskussion

Bei der Planung eines Versuchsdesigns lassen sich drei bedeutsame Phasen unterscheiden. Dies ist zum einen die Operationalisierung, also die Umsetzung der Fragestellungen in messbare und somit

nachvollziehbare Dimensionen. Ist dies geschehen, so muss ein geeignetes Versuchsdesign entwickelt werden. Dieses sollte geeignet sein, die Fragestellungen zu beantworten, was unter anderem mit der Wahl geeigneter Versuchs- und Kontrollgruppen einhergeht. Im letzten Schritt müssen Überlegungen zur Kontrolle möglicher Störvariablen getroffen werden (Kühl et al. 2009). Wie diese Punkte im Rahmen der vorliegenden Evaluierung Berücksichtigung finden, soll im Folgenden diskutiert werden.

4.2.1.1 Operationalisierung

Zur Beantwortung der unter Punkt 1.2.5 aufgeführten Fragestellungen, war es notwendig, geeignete kognitive und affektive Messinstrumente zu finden, die die gewünschten Konstrukte erfassen. Hier konnte weitgehend auf bewährte Messinstrumente zurückgegriffen werden, welche in den letzten Jahren in anderen Schülerlaborstudien entwickelt und optimiert wurden (z.B. bei Engeln 2004, Brandt 2005, Pawek 2009). Wie bereits erörtert wurde, ist deren Objektivität, Reliabilität und Validität anzunehmen. Wie der im Rahmen der vorliegenden Arbeit neu entwickelte FSWEx diese Gütekriterien erfüllt, wird der folgende Abschnitt zeigen. Hier wird auch unmittelbar die erste der Fragestellungen beantwortet:

Lassen sich die drei Teildimensionen des Modells experimenteller Kompetenz nach Schreiber et al. (2009) in Likert-skalierten Items operationalisieren und in ein reliables Messinstrument zur Erfassung der experimentbezogenen Selbstwirksamkeit überführen?

Anlass für die Entwicklung des FSWEx waren die als wenig valide kritisierten schriftlichen Verfahren (Tannenbaum 1971; Fraser 1980), sowie die wenig ökonomischen praktischen Testverfahren zur Erfassung der Experimentierkompetenz (Tamir 1974, Ben-Zvi et al. 1977, Schreiber et al. 2009). Der FSWEx geht daher einen anderen Weg. Er erhebt zunächst nicht den Anspruch, die praktische Experimentierfähigkeit direkt zu erfassen, sondern misst die subjektiv eingeschätzte Experimentierfähigkeit einer Person, hier auch als experimentbezogene Selbstwirksamkeitserwartung bezeichnet, in den Subskalen Planung, Durchführung und Auswertung. Die Itemformulierung erfolgte theoriebasiert auf Grundlage des Modells experimenteller Kompetenz nach Schreiber et al. (2009). Als Itemformat wurde in Anlehnung an ähnliche Messinstrumente (Wilde et al. 2009) eine fünfstufige Likert-Skala gewählt. Diese ist sicher als objektiv auswertbar zu betrachten. Die Reliabilität der im Rahmen der Hauptstudie entwickelten Fragebogenversion FSWEx.V1 erwies sich mit Cronbachs α -Werten zwischen 0,513 und 0,647 zwar im Sinne von Schön (2007) und Ozkan et al. (2008) als zufriedenstellend, allerdings als noch verbesserungswürdig. In Anlehnung an das Modell experimenteller Kompetenz wurden weitere Items in Form des FSWEx.V2

entwickelt, welche eine weitergehende Abdeckung des Modells versprochen. Mit Cronbachs α -Werten zwischen 0,729 und 0,789 konnten gute Skalenreliabilitäten erzielt werden.

Weiterhin gelang es faktorenanalytisch sowie über Interkorrelationen für beide eben genannten FSWEx-Versionen zu belegen, dass die entwickelten Items tatsächlich die drei unterschiedlichen Skalen abdecken. Lediglich einzelne Items wiesen Streuladungen, beziehungsweise eine Hauptladung auf einem anderen, als den erwarteten Faktoren auf. Als Ladung bezeichnet man die Korrelation zwischen einer Variable und einem Faktor. Bei hohen Nebenladungen (also auf anderen als den erwarteten Faktoren) ist die Zuordnung einer Variable zu einem Faktor uneindeutig. Eine Faktorenanalyse dient dazu, aus den hier aufgenommenen Daten mehrerer verschiedener manifester Variablen, also Items, auf wenige zugrunde liegende latente Variablen (also Faktoren) zu schließen. In der vorliegenden Untersuchung wurde das Verfahren der explorativen Faktorenanalyse realisiert, da sie typischerweise bei der Entwicklung eines ganz neuen Messinstruments eingesetzt wird (vgl. Abschnitt 2.2.5.9). Sicherlich wäre im vorliegenden Fall auch alternativ die Durchführung einer confirmatorischen Faktorenanalyse gerechtfertigt gewesen, da der FSWEx zwar neu, aber anhand eines theoretisch etablierten Modells (Schreiber et al. 2009), mit einer begründeten faktoriellen Struktur, formuliert wurde.

In beiden FSWEx-Versionen fallen insbesondere die Items „Es gelingt mir, anhand theoretischer Überlegungen Erwartungen über den Ausgang eines Experimentes zu formulieren“ und „Mit dem Aufschreiben von Versuchsbeobachtungen habe ich Schwierigkeiten“ auf. Das erste Item wurde konstruiert zur Erfassung der Modellaspekte *Erwartungen formulieren* und *Hypothesen bilden*, welche im Modell nach Schreiber et al. (2009) der Subskala Planung zugeordnet werden. Das zweite Item operationalisiert den Modellaspekt *Messungen dokumentieren*, welcher im Modell der Subskala Durchführung angehört. Beide Items laden allerdings vorwiegend auf den Faktor Auswertung, d.h. sie korrelieren mit diesem. Im Rahmen der vorliegenden Studie verblieben die Items in der geplanten Subskala, sofern dies zu keiner deutlichen Reliabilitätsverschlechterung führte. Dies gewährleistete eine bessere Modellabdeckung. Bevor für zukünftige Studien eine Neuordnung oder Umformulierung der Items in Erwägung gezogen wird, bleibt zu berücksichtigen, dass die Subskala Durchführung im Modell noch die Aspekte *Geräte zusammenstellen* und *Versuchsanordnungen aufbauen* enthält. Diese wurden bewusst in den bisherigen FSWEx noch nicht aufgenommen, da das Messinstrument in der vorliegenden Studie konkret im Zusammenhang mit Laborkursen eingesetzt werden sollte, die keine praktischen Tätigkeiten im Sinne der beiden letzten Modellaspekte beinhalteten. Die Aufnahme von entsprechenden Items in die Subskala Durchführung könnte dazu führen, dass das Item „Mit dem Aufschreiben von Versuchsbeobachtungen habe ich Schwierigkeiten.“ besser gemeinsam mit den Items der Subskala Durchführung auf einen Faktor lädt.

Für das Item „Es gelingt mir, anhand theoretischer Überlegungen Erwartungen über den Ausgang eines Experimentes zu formulieren“ erscheint inhaltlich keine Zuordnung zur Subskala Auswertung sinnvoll. In Bezug auf dieses Item sollte in Erwägung gezogen werden, ob die drei Subskalen des Modells das Konstrukt der Experimentierkompetenz hinreichend differenziert abbilden, oder ob eine weitere Unterteilung der Subskala Planung sinnvoll erscheint. So differenziert Mayer (2007) explizit in seinem Modell naturwissenschaftlichen Denkens zwischen den Teildimensionen *Naturwissenschaftliche Fragen formulieren*, *Hypothesen generieren* und *Untersuchungen planen*. Diese spiegeln inhaltlich relativ gut die Unterpunkte der Teilkompetenz Planung im Modell nach Schreiber et al. (2009) wieder. Auch unabhängig von der in Modellen beschriebenen Experimentierkompetenz wird der Prozess des Experimentierens häufig in weitere Teilschritte untergliedert. So sieht Brandt (2005) Experimente als Aspekt des forschenden Lernens. In ihnen werden im gesamten Forschungsprozess vom Phänomen über die Hypothesenbildung, Experimentplanung und -durchführung, Interpretation und Dokumentation der Experimente alle Anteile bearbeitet. Eine Ausdifferenzierung des FSWEx um weitere Items erscheint somit sinnvoll, um eine vollständigere und differenziertere Abdeckung des Konstruktes der Experimentierkompetenz zu gewährleisten. Die Frage der tatsächlichen Dimensionalität der Experimentierkompetenz kann und soll hier also nicht geklärt werden.

Nun wurden im Rahmen der vorliegenden Untersuchung nicht nur die eben erwähnten Messinstrumente FSWEx.V1 und V2 entwickelt, welche die experimentbezogene Selbstwirksamkeit unabhängig von einer konkreten Experimentiersituation erfassen. Im Rahmen der Hauptstudie wurde zudem der postFSWEx.V1 eingesetzt, um unmittelbar im Anschluss an eine Experimentiersituation die situationsbezogene Selbstwirksamkeit zu erfassen. So zeigte dieses Messinstrument mit Reliabilitäten zwischen Cronbachs $\alpha=0,614$ und $0,768$ trotz seiner Kürze zufriedenstellende Skalenreliabilitäten. Im Rahmen der Faktorenanalyse wurde lediglich das Item „Mit dem Aufschreiben von Versuchsbeobachtungen habe ich Schwierigkeiten.“ als auffällig bewertet, da es nicht wie erwartet gemeinsam mit den anderen Items der Subskala Durchführung auf einen Faktor lud, sondern mit den Items der Subskala Auswertung. Dies verwundert an dieser Stelle nicht, denn ein Item fast identischer Formulierung wurde bereits im FSWEx.V1 und V2 als problematisch erkannt. Wie dies bereits für die anderen beiden Versionen des Messinstrumentes angemerkt wurde, so wurde auch der postFSWEx darauf hin entworfen, die subjektiv empfundene Selbstwirksamkeit in einer Experimentiersituation zu erfassen, die schwerpunktmäßig das praktische Durchführen und Auswerten von Versuchsergebnissen umfasste. Sollte er eingesetzt werden, um unmittelbar nach einer Experimentiersituation die erlebte Planungskompetenz zu erfassen, so sollten die Items der entsprechenden Subskala umformuliert werden. Ferner bietet sich auch beim postFSWEx.V2 die

Aufnahme weiterer Items für eine differenziertere Erfassung des Konstruktes der subjektiv eingeschätzten Experimentierkompetenz an.

Unmittelbar an die vorangegangenen Überlegungen stellt sich die Frage der Validität des Messinstrumentes FSWEx. Ein valides Modell experimenteller Kompetenz angenommen, sollte auch der FSWEx eine hohe Inhaltsvalidität besitzen. Hier muss einschränkend betont werden, dass die FSWEx-Items – wie im vorherigen Abschnitt eingehend erläutert – noch nicht beabsichtigen, das vollständige Modell experimenteller Kompetenz abzubilden. Eine Aufnahme ergänzender Items, insbesondere für die Subskala Durchführung würde somit potentiell zu einer Optimierung der Inhaltsvalidität führen. Zur Beurteilung der Kriteriumsvalidität des FSWEx wurden für die Subskalen aller FSWEx-Versionen Korrelationen mit dem Außenkriterien des biologischen Fähigkeitsselbstkonzeptes berechnet, da Selbstkonzept und Selbstwirksamkeit theoretisch in engem Verhältnis stehen. Sämtliche Korrelationen (r = Korrelationskoeffizient) lagen im Bereich zwischen $r=0,336$ und $r=0,554$. Lediglich die Korrelation des Fähigkeitsselbstkonzeptes der FSWEx.V1 Subskala Planung fiel mit $r=0,165$ deutlich geringer aus, was auf eine nur geringe Korrelation zwischen dem Selbstkonzept und dieser Selbstwirksamkeitsskala schließen lässt. Fraglich ist, wie die Höhe der Korrelationen zur Beurteilung der Kriteriumsvalidität einzuschätzen ist. In ihrer Studie zur Entwicklung eines Messinstrumentes zur Erfassung der Selbstwirksamkeit zu sportlicher Aktivität korrelieren Fuchs & Schwarzer (1994) ihr Messinstrument unter anderem mit einer Skala, welche die Intention zum Sport treiben erfasste und deuteten eine Korrelation von $r=0,47$ als einen Hinweis für dessen Validität. In ähnlicher Höhe liegen bis auf die Korrelation der FSWEx.V1 Subskala Planung auch die hier berichteten Korrelationen. Dies mag nicht sonderlich verwundern, da hier das biologische Fähigkeitsselbstkonzept mit der Selbstwirksamkeitserwartung am Experimentieren korreliert wurde. Zwar wird durchaus ein positiver Zusammenhang angenommen, allerdings ist nicht zu erwarten, dass jede Person, die sich im Fach Biologie kompetent einschätzt, sich auch gleichzeitig als guten Experimentierer wahrnimmt. Das Selbstkonzept und die Selbstwirksamkeit einer Person lassen sich zu zahlreichen Subkonstrukten ausdifferenzieren (Möller & Trautwein 2009, Moschner & Dickhäuser 2010). Entsprechend belegen verschiedene Studien, dass die Korrelationen zwischen Selbstkonzept und Selbstwirksamkeit, beziehungsweise innerhalb verschiedener Facetten des Selbstkonzeptes oder der Selbstwirksamkeit mitunter deutlich variieren (Fuchs & Schwarzer 1994; Schulte 2008; Möller & Trautwein 2009). Somit sprechen die Korrelationen der FSWEx-Subskalen mit dem biologischen Fähigkeitsselbstkonzept sicher nicht gegen deren Validität. Insbesondere in Anbetracht der Tatsache, dass die Korrelationen der drei Subskalen untereinander in einer ähnlichen Größenordnung liegen, macht deutlich, dass die Korrelationen zum biologischen Selbstkonzept sogar relativ hoch ausfallen.

Zwar ist die Erhebung der Selbstwirksamkeit an sich schon von besonderem Interesse, da sie im unmittelbaren Zusammenhang mit Leistungsmotivation und Verhalten gesehen wird. Doch ist fraglich, wie die experimentbezogene Selbstwirksamkeitserwartung tatsächlich die Experimentierkompetenz einer Person widerspiegelt. Eine entsprechende Fragestellung wurde bereits unter 1.2.5 formuliert und soll nachfolgend anhand der vorliegenden Ergebnisse erörtert werden.

Liefert die experimentbezogene Selbstwirksamkeitserwartung einer Person Hinweise auf ihre tatsächliche Kompetenz beim Experimentieren?

Für die Beantwortung dieser Fragestellung wurden die Fragebogendaten der FSWEx.V1- und postFSWEx.V1-Subskala Durchführung, welche im Laborkurs Genetischer Fingerabdruck der Hauptstudie erhoben wurden, mit den von Dahmen (2012) im Rahmen einer Videoanalyse aufgezeichneten Daten korreliert. Hier stellten sich lediglich wenige und eher gering signifikante und tendenzielle Korrelationen heraus, die allerdings in der Richtung ihres Vorzeichens weitgehend den Erwartungen entsprachen. Schüler mit höherer Selbsteinschätzung zeigten weniger Fehler in der Bedienung von Geräten, sie handelten eher experimentbezogen und zeigten sich weniger unsicher. Lediglich Schüler, die im Follow-up-Test ihre Durchführungskompetenz hoch einschätzten, zeigten bei der Pipettierübung ein unerwartetes Verhalten, da sie häufiger nicht experimentorientiert kommunizierten. Dieser Zusammenhang ist mit einer Irrtumswahrscheinlichkeit von $p=0,060$ allerdings nur tendenziell gegeben.

Zwei Ursachen sind zur Erklärung der beobachteten geringen Korrelationen denkbar. Zum einen besteht die Möglichkeit, dass Schüler ihre Experimentierkompetenz nur mäßig treffend einschätzen, zum anderen könnte die Umsetzung der Videoanalyse Auswirkungen auf die Einschätzung der Experimentierkompetenz haben. Der Fokus der Videoanalyse wurde von Dahmen (2012) bewusst auf zwei relativ kurze Experimentiersituationen gerichtet, um möglichst viele Schüler bei der gleichen Experimentiertätigkeit erfassen zu können. Vorteilhaft ist, dass eine ausreichende Stichprobengröße für eine sinnvolle quantitative Auswertung erzielt wurde. Als nachteilhaft könnte sich herausstellen, dass die etwa einminütigen Videosequenzen nur einen kurzen Teil der Realität abbilden (Dinkelaker & Herrle 2009; Dahmen 2012). Es ist somit nicht auszuschließen, dass die dokumentierten Experimentiertätigkeiten nicht hinreichend repräsentativ zur Abbildung der praktischen Durchführungskompetenz sind. Insbesondere die dokumentierte Pipettierübung fand unmittelbar nachdem die Schüler den Laborraum betreten hatten statt. Dies führte – bedingt durch den Neuigkeitscharakter der Umgebung sowie der zu verwendenden Sicherheitsausrüstung (Kittel, Schutzbrille, Handschuhe) – zunächst zu einer regen Kommunikation der Schülergruppen. Diese könnte mit ursächlich dafür sein, dass bei der Pipettierübung die sich kompetenter einschätzenden

Schüler häufiger nicht experimentbezogen kommunizierten. Ihre Sicherheit im Umgang mit den Pipetten (diese wird durch die geringen Korrelationen mit der FSWEx-Subskala Durchführung gestützt) könnte Aufmerksamkeitsreserven für die experimentfremde Kommunikation entsprechender Schüler eröffnet haben. Die ansonsten schlüssigen Korrelationen deuten einen positiven Zusammenhang der FSWEx-Subskala Durchführung und dem realen Experimentierverhalten an, so dass eine gänzliche Fehleinschätzung der Schüler unwahrscheinlich erscheint.

Allerdings stellt sich die Frage, wie groß der optimistische Anteil bei der Einschätzung der Handlungskompetenz ist. So könnten Selbstwirksamkeitserwartung bei einigen Schülern Ausdruck von Über- oder Unterschätzung sein und bei anderen Schülern auf Erfolgserfahrungen beruhen (Satow 2000). Eine starke Abhängigkeit der Selbsteinschätzung wird auch dem kulturellen Bezugssystem beigemessen. So erreichen bei PISA 2006 beispielsweise Japan und Korea hohe Kompetenzwerte, allerdings gaben die Schüler in Bezug auf ihre Selbsteinschätzung nur niedrige Kennwerte an (Prenzel 2007). Zehren (2009) äußert zudem die Vermutung, dass angeleitetes Arbeiten eine realistische Selbsteinschätzung erschwert. Fraglich ist also, ob die Schüler aus dem schulischen Kontext überhaupt ausreichend praktische Erfahrung mitbringen, um ihre Experimentierfähigkeit realistisch einschätzen zu können. In kommenden Untersuchungen wird es daher von Interesse sein, längere Videosequenzen zu wählen, welche einen ganzheitlicheren Eindruck der Experimentierkompetenz vermitteln. Korrelationen mit der Selbstwirksamkeitserwartung sollten im Falle realistischer Selbsteinschätzungen der Schüler dann zu deutlicheren Korrelationen mit dem FSWEx führen. Weiterhin konnten bisher lediglich – bedingt durch den Fokus der Videoaufzeichnungen von Dahmen (2012) – Korrelationen zwischen dem Experimentierverhalten und der FSWEx-Subskala Durchführung vorgenommen werden. Zukünftig wird es interessant sein, auch die Subskalen Planung und Auswertung mit dem entsprechenden realen Verhalten abzugleichen. Insbesondere für den Prozess der Planung, welcher kaum in Form von Videos zu dokumentieren ist, wird dazu die Entwicklung weiterer Messinstrumente notwendig sein.

Es bleibt festzuhalten, dass mit dem FSWEx ein ökonomisches Messinstrument entwickelt wurde, welches in drei Subskalen die experimentbezogene Selbstwirksamkeit (postFSWEx.V1) bzw. die experimentbezogene Selbstwirksamkeitserwartung (FSWEx.V1 & V2) erfragt. Die Bearbeitungszeit für Schüler, aber auch für Lehrkräfte, beträgt für den vollständigen FSWEx.V2 weniger als zehn Minuten. Für zukünftige Studien empfiehlt sich insbesondere die Verwendung des FSWEx.V2, da er eine gute Reliabilität sowie eine umfassendere Abdeckung des Modells nach Schreiber et al. (2009) gewährleistet. In kommenden Studien gilt es, die Validität des Messinstrumentes durch weitere Außenkriterien zu bestätigen. Insbesondere umfassende Videoanalysen können dazu beitragen, den

Zusammenhang von Selbsteinschätzung und tatsächlicher Experimentierkompetenz weiter aufzuklären.

Während vorangegangen die Operationalisierung eines für die vorliegende Studie zentralen Konstruktes diskutiert wurde, sollen im folgenden Abschnitt einige Aspekte des Versuchsdesigns erörtert werden.

4.2.1.2 Versuchsdesign

Wie bereits vorgestellt, wurde die vorliegende Haupt- und Nebenstudie – wie in der Lehr-Lernforschung üblich – als quasiexperimentelle Untersuchung umgesetzt. Bei den Stichproben handelt es sich somit nicht um echte Zufallsstichproben, wie sie unter experimentellen Bedingungen zu fordern sind, sondern um natürlich gewachsene Gruppierungen mit ihren jeweils spezifischen Besonderheiten. In einer quasiexperimentellen Studie ist es somit kaum vermeidbar, dass sich die zu vergleichenden Gruppen in verschiedenen Variablen unterscheiden. Ergeben sich zwischen den zu vergleichenden Gruppen Unterschiede in Bezug auf abhängige Variablen, so können diese nicht eindeutig auf eine unabhängige Variable zurückgeführt werden. Eine quasiexperimentelle Untersuchung verfügt daher im Vergleich zu einer experimentellen Untersuchung über eine geringere interne Validität (Bortz & Döring 2006). Eine experimentelle Untersuchung wäre in diesem Rahmen nur möglich, wenn die Gruppen der Schüler, welche an den Laborkursen des BeLL Bio teilnehmen, nach dem Zufall zusammengesetzt würden. Bei einer ausreichend großen Stichprobe würden sich etwaige Unterschiede zwischen den Gruppen herausmitteln. Zwar würde diese Zusammensetzung zu einer höheren internen Validität führen, jedoch würde jegliche Art sozialer Interaktionen, wie sie eine natürlich gewachsene Schulklasse determinieren, unberücksichtigt bleiben. Es ist kaum davon auszugehen, dass sich zufällig „zusammengewürfelte“ Schüler so verhalten, wie sie dies im Rahmen ihres gewohnten Klassenverbandes tun würden. Bei einem derartigen Design erscheint es somit fraglich, ob die so erzielten Ergebnisse auch auf natürliche Situationen generalisierbar wären. Entsprechend würde es der streng kontrollierten Studie an externer Validität mangeln (Bortz & Döring 2006). Die vorliegende, im Kontext des BeLL Bio durchgeführte Studie, soll Aufschluss geben über die Wirkung der angebotenen Laborkurse auf Schulklassen bzw. Kurse der gymnasialen Oberstufe. Diese natürlichen Gruppen sind die Zielgruppe des BeLL Bio. Die Ergebnisse einer quasiexperimentellen Studie, in welcher sich aus genau diesen Gruppen die Stichprobe rekrutiert, sichert somit eine hohe externe Validität und eine hohe Transferfähigkeit der Ergebnisse auf zukünftige, das BeLL Bio besuchende Schülerkurse. Zudem erscheint die untersuchte Stichprobe sehr repräsentativ für die das BeLL Bio besuchenden

Oberstufenkurse, da im Datenerhebungszeitraum fast jede das Labor besuchende Schulklasse in der Evaluierung berücksichtigt werden konnte.

Sicherlich sind diese Ergebnisse nicht ohne weiteres auf Schülerlabore übertragbar, in denen Schüler nicht im Klassenverband, sondern interessenabhängig das Labor besuchen. Auch werden die Ergebnisse wenig über die Wirkung von konzeptionell offener gestalteten Schülerlaboren aussagen. Ein Blick in die Schülerlaborszene zeigt aber, dass dem BeLL Bio ähnliche Kurskonzepte in vielen naturwissenschaftlichen Schülerlaboren Anwendung finden. In vielen Fällen besuchen die Schüler im Klassenverband und nicht gemeinsam mit ihnen unbekannten Schülern ein Labor. Häufig dienen Versuchsskripte als Anhalt für die durchgeführten Versuche. Die im Rahmen der vorliegenden Studie erhaltenen Ergebnisse sind damit in hohem Maße übertragbar auf viele andere naturwissenschaftliche Schülerlabore, so dass sie einen wichtigen Beitrag zur noch relativ jungen Schülerlaborforschung liefern.

Auch wenn eine hohe externe Validität der vorliegenden Studie gesichert ist, so bleibt in der vorliegenden Studie auch das Problem der mangelnden internen Validität quasiexperimenteller Studien nicht unberücksichtigt. Diesem Problem wurde durch Einrichtung einer Kontrollgruppe, sowie im Rahmen der Datenauswertung Rechnung getragen.

4.2.1.3 Kontrolle von Störvariablen

Die vorliegende Evaluierung wurde durchgeführt, um die kognitive und affeektive Wirksamkeit eines Laborbesuches im BeLL Bio belegen zu können. Dazu gehört auch der Nachweis, dass die registrierten Effekte ohne Einsatz des Laborbesuches ausbleiben. Nur so kann sichergestellt werden, dass tatsächlich der Laborbesuch und keine andere Einflussgröße das beobachtete Ergebnis verursacht (Bortz & Döring 2006). Der Wahl einer geeigneten Kontrollgruppe wird für die Untersuchung von Lehr- und Lernsituationen daher hohe Bedeutung beigemessen (Randler & Bogner 2008). Häufig fällt diese Wahl gerade im Rahmen quasiexperimenteller Studien schwer, da hier keine echten Zufallsstichproben zu realisieren und damit auch kaum eine der Vergleichsgruppe ähnliche Kontrollgruppe zu erreichen ist. Aus diesem Grund verzichten beispielsweise (Engeln 2004, Glowinski 2007) in ihren Schülerlaborevaluierungen gänzlich auf Kontrollgruppen. Glowinski (2007) argumentiert, dass ein Schülerlabor als multifaktoriell wirksame Lernumgebung verstanden werden muss. An dieser wirkten verschiedenste Einflussfaktoren, welche in ihrer Einzelwirkung nicht kontrollierbar seien, so dass als Kontrollgruppe kaum eine einzelne Gruppe definiert werden könne. Glowinski (2007) geht somit davon aus, dass Erkenntnisse aus einer entsprechenden Kontrollgruppe

mangels externer Validität für eine Evaluation von Schülerlaboren als Lernumgebung wenig hilfreich erscheinen.

Auch wenn die Problematik der Wahl einer geeigneten Kontrollgruppe offensichtlich ist, so wird sie in der vorliegenden Studie doch als unabdingbar betrachtet. Insbesondere im Rahmen der Längsschnittanalysen erfüllt die Kontrollgruppe die Funktion, Pretesteffekte, also Veränderungen, welche allein durch das Ausfüllen des Fragebogens bewirkt werden, auszuschließen (Groß & Großengießer 2003). Darüber hinaus findet die Kontrollgruppe in der vorliegenden Studie Verwendung, um ein Bezugsmaß für die Höhe der Ausprägung der Wahrnehmung des Lernortes Labor zu haben. Auch Bortz und Döring (2006) betonen, dass eine Kontrollgruppe unabdingbar ist, um einen auf eine Maßnahme zurückgehenden Effekt in der Richtigen Größenordnung erfassen zu können. In der vorliegenden Studie stellte sich für die Wahl der Versuchsgruppe als besonders problematisch heraus, dass hier nicht nur affektive, sondern auch kognitive Variablen erhoben werden sollten. Dies bedeutete, dass die Kontrollgruppe im Untersuchungszeitraum das Laborkursthema nicht im schulischen Biologieunterricht behandeln durfte. Anderenfalls wäre nicht nachvollziehbar gewesen, ob ein eventueller Lernzuwachs der Kontrollgruppe durch einen Pretesteffekt oder durch den schulischen Biologieunterricht bewirkt wurde. Da Schulklassen das BeLL Bio in der Regel in dem Schuljahr besuchen, in welchem das Laborkursthema im Biologieunterricht behandelt wird, musste für die Zusammensetzung der Kontrollgruppe in einigen Fällen auf die darüber- oder darunterliegende Jahrgangsstufe ausgewichen werden. Als ebenfalls problematisch erwies sich, dass viele der angesprochenen Lehrkräfte – einmal auf das BeLL Bio aufmerksam gemacht – gerne mit ihrem Biologiekurs das Labor besuchen, aber nicht als Kontrollgruppe ohne Laborbesuch fungieren wollten. Aus diesem Grund wurde in der vorliegenden Studie für alle Versuchsgruppen eine gemeinsame Kontrollgruppe eingerichtet, die sich bemüht, die Versuchsgruppen in ihren dispositionalen Merkmalen möglichst gut zu repräsentieren.

Um dies zu kontrollieren wurden bereits im Pretest verschiedene dispositionale Variablen erfasst, um diese auf Unterschiede zwischen der Versuchs- und Kontrollgruppen zu prüfen. Hier wurde als bedeutsamster Unterschied ein höheres Fachinteresse der Kontrollgruppenschüler gegenüber den Versuchsgruppenschülern festgestellt. Auch wenn die geringe Effektstärke einen nur mäßigen Einfluss der Gruppenzugehörigkeit auf das Fachinteresse belegte (vgl. Cohen 1988), wurde das dispositionale Fachinteresse in den Datenauswertungen als Kovariate berücksichtigt, in denen ein bedeutsamer verzerrender Einfluss auf abhängige Variablen anzunehmen war. Der Einfluss der Kovariaten wurde in diesem Fall aus der abhängigen Variable herausgerechnet. Unter Punkt 1.2.3.4 wird berichtet, dass das dispositionale, neben dem situationalen Interesse Einfluss auf das aktuelle Interesse hat (Krapp 2010, Mitchell 1993). Im Rahmen der Längsschnittanalysen wird das Interesse

der Kontrollgruppe am Experimentieren im Biologieunterricht in Relation gesetzt zu dem aktuellen Interesse, welches die Schüler der Versuchsgruppe unmittelbar im Anschluss an die Laborkurse empfinden. Da entsprechend der Interessentheorie ein bedeutsamer Einfluss des Fachinteresses auf dieses aktuelle Interesse anzunehmen ist, wurde der Einfluss des Fachinteresses als Kovariate in die entsprechenden Analysen des aktuellen Interesses einbezogen. Im Rahmen der Längsschnittanalysen wurden zudem sämtliche Vergleichsgruppen auf Pretestäquivalenz untersucht. Dies galt sowohl für die Vergleiche zwischen Versuchs- und Kontrollgruppe, als auch für alle nach Geschlecht, Migrationshintergrund, beruflicher Orientierung und Schülertypus durchgeführten Subgruppenanalysen innerhalb der einzelnen Versuchsgruppen. Lagen bereits im Pretest bedeutsame Differenzen vor, so wurde – in Anlehnung an andere quasiexperimentelle Studien – die Pretestmessungen als Kovariate berücksichtigt (Arisholm et al. 2007; vgl. Natzke & Petermann 2009). Dieses Vorgehen diene dazu, die durch das quasiexperimentelle Design bedingten Ungleichgewichte zwischen den Vergleichsgruppen auszugleichen.

4.2.2 Ergebnisdiskussion

Im Rahmen der Methodendiskussion wurde bereits auf die ersten beiden der unter Punkt 1.2.5 formulierten Fragestellungen eingegangen. Entsprechend orientiert sich die folgende Ergebnisdiskussion an den weiteren Fragen. Viele der Befunde ähneln sich für die drei Versuchsgruppen, so dass die Ergebnisdiskussion gemeinsam erfolgen kann. Auf bedeutsame, versuchsgruppenspezifische Beobachtungen wird explizit hingewiesen.

4.2.2.1 Experimentbezogene Selbstwirksamkeit

Aus der Überlegung heraus, dass der Selbstwirksamkeit wesentliche Bedeutung im Zusammenhang mit der Entwicklung von Leistungsmotivation und Anstrengungsbereitschaft beigemessen wird (Woolfolk 2008), ergab sich folgende Überlegung.

Fraglich ist, wie Schüler ihre Experimentierkompetenz in praktischen Experimentierkursen im Schülerlabor einstufen und wie sich der Laborbesuch auf die experimentbezogene Selbstwirksamkeitserwartung auswirkt.

Wie im Rahmen von Haupt- und Nebenstudie gezeigt werden konnte, schätzen Schüler ihre Planungs-, Durchführungs- und Auswertungskompetenz in der Höhe sehr unterschiedlich, aber nach einem identischen Einschätzungsschema ein. Durchweg findet sich – unabhängig von Messzeitpunkt, Geschlecht, Migrationshintergrund, beruflicher Orientierung oder Schülertypus– die deutliche geringste Kompetenzeinschätzung im Bereich der Planungskompetenz. Deutlich höher fällt in der Regel die Selbstwirksamkeitserwartung aus, Experimente erfolgreich auswerten zu können. Noch

höher schätzen Schüler ihre Kompetenz ein, Experimente erfolgreich durchführen zu können. Diese Beobachtung verwundert in Anbetracht der Resultate der PISA 2006-Erhebung nicht. Hier wurde festgestellt, dass der Forschungsaspekt beim Experimentieren in der Schule deutlich zu kurz kommt. Zwar werden Experimente regelmäßig im naturwissenschaftlichen Unterricht durchgeführt, allerdings liegt in Deutschland der Schwerpunkt auf Demonstrationsexperimenten. Die Möglichkeit im Unterricht eigene Fragestellungen zu entwickeln und naturwissenschaftliche Untersuchungen zu planen, ist in Deutschland im internationalen Vergleich nur sehr gering ausgeprägt. Besonders selten werden Experimente selbst entwickelt (Seidel et al. 2007). Entsprechend scheinen die Einschätzungen der Schüler in der vorliegenden Studie realistisch, wenn sie ihre Planungskompetenz eher gering einstufen.

Wie sich in der Hauptstudie zeigte, ist die Heterogenität der Schüler bezüglich der Höhe ihrer Kompetenzeinschätzungen allerdings sehr groß. Während sich keine bedeutenden Abhängigkeiten der Selbstwirksamkeitserwartung vom Migrationshintergrund beobachten lassen, schätzen männliche Schüler sowohl in der Hauptstudie, als auch in der Nebenstudie ihre Planungskompetenz geringfügig höher ein, als die weiblichen Schüler. Allerdings erwies sich dieser Unterschied nur im Pretest des Kurses Genetischer Fingerabdruck als statistisch bedeutsam. Diese geringere Kompetenzeinschätzung bei den weiblichen Schülern ist wahrscheinlich weniger auf tatsächliche Leistungsunterschiede zwischen den Geschlechtern, sondern vermutlich eher auf eine Tendenz zur Überschätzung der männlichen Schüler zurückzuführen. Nennenswerte Unterschiede in der naturwissenschaftlichen Kompetenz zwischen männlichen und weiblichen Schülern konnten bei PISA 2006 nämlich nicht festgestellt werden (Prenzel 2007). Die Überschätzung männlicher Schüler bezogen auf ihre schulische Leistungsfähigkeit wird dagegen häufiger berichtet (Helbig 2012).

Deutlichere Unterschiede sind in Abhängigkeit von der beruflichen Orientierung und dem Schülertypus festzustellen. Naturwissenschaftlich orientierte Schülerinnen und Schüler bekunden zu allen Messzeitpunkten eine höhere Planungs-, Durchführungs- und Auswertungskompetenz, als die nicht naturwissenschaftlich Orientierten. Ähnliche Unterschiede finden sich auch bei der Differenzierung nach den in der vorliegenden Studie identifizierten Schülertypen. Zu allen Messzeitpunkten zeichnet sich der Allrounder durch die höchsten Selbsteinschätzungen in allen FSWE-Subskalen aus. In der Regel schätzt dieser Schülertyp seine Experimentierfähigkeit deutlich höher ein, als die Praktiker und Theoretiker. Praktiker und Theoretiker liegen in der Einschätzung ihrer Durchführungs- und Planungskompetenz auf ähnlichem Niveau. Ihre Auswertungskompetenz schätzen hingegen die Theoretiker in der Regel höher ein, als die Praktiker. Die Abhängigkeit der selbst eingeschätzten Kompetenz von Berufsorientierung und Schülertyp verwundert nicht, denn sowohl bei den naturwissenschaftlich orientierten Schülern, als auch bei den Allroundern handelt es

sich um fachlich und sachlich sehr interessierte Schüler mit hohem biologischen Fähigkeitsselbstkonzept. Die experimentbezogene Selbstwirksamkeit ist als Bestandteil des globaleren fachlichen Fähigkeitskonzeptes anzusehen (Köller & Möller 2010). Entsprechend fallend die Kompetenzerwartungen bei den Schülern des geringeren fachlichen Selbstkonzeptes auch eher geringer aus. Allerdings wird hier deutlich, dass es sich bei den Teildimensionen der experimentbezogenen Selbstwirksamkeitserwartung tatsächlich um unterschiedliche Teilkonstrukte handelt. Ansonsten wäre anzunehmen, dass der Theoretiker, welcher ebenfalls ähnlich den Allroundern ein relativ hohes fachliches Selbstkonzept aufweist, auch in der Durchführungskompetenz höhere Einschätzungen bekundet, als der Praktiker, denn dieser hat ein deutlich geringeres fachliches Selbstkonzept. Offensichtlich spielen für die empfundene Durchführungskompetenz noch andere Aspekte eine Rolle, als das fachliche Selbstkonzept. Denkbar ist, dass der Praktiker sich im Sinne internaler sozialer Vergleiche (Pohlmann & Möller 2006) in früheren Experimentiersituationen im Vergleich zu seinen Mitschülern als kompetent empfunden hat und diese Erfahrung zur aktuellen Einschätzung seiner Durchführungskompetenz beiträgt.

Doch hat der Besuch im Schülerlabor BeLL Bio Auswirkungen auf die subjektiv eingeschätzte Experimentierkompetenz? Hier wurde festgestellt, dass sich die experimentbezogene Selbstwirksamkeitserwartung, ähnlich wie das biologische Fähigkeitsselbstkonzept, als zeitlich langfristig stabil herausstellte. Eine bedeutsame kurzfristige Steigerung der experimentbezogenen Selbstwirksamkeit in der Subskala Durchführung konnte allerdings in der Nebenstudie berichtet werden. Hier zeigte sich, dass der Besuch des Laborkurses Genetischer Fingerabdruck unabhängig vom Geschlecht der Schüler zu einer deutlichen Steigerung der subjektiv empfundenen Durchführungskompetenz führte. Diese ist auch Wochen nach dem Laborbesuch noch tendenziell nachzuweisen. In der Subskala Planung konnte zudem bei den männlichen Schülern ein gegenüber den weiblichen Schülern deutlicher, kurzfristiger Anstieg der empfundenen Planungskompetenz beobachtet werden. Dieser erwies sich im Vergleich zur Kontrollgruppe allerdings als nicht bedeutsam. Eine zeitliche Änderung in der wahrgenommenen Auswertungskompetenz ist weder kurz- noch langfristig statistisch zu belegen.

Nun ist die Fragestellung von Interesse, wie der Laborbesuch überhaupt eine Veränderung der Selbstwirksamkeitserwartung herbeiführten konnte und warum keine durch den Laborbesuch bedingten Veränderungen in den Subskalen Planung und Auswertung festgestellt wurden. Da in der Nebenstudie lediglich die experimentbezogene Selbstwirksamkeit erhoben wurde, muss von den Ergebnissen der Hauptstudie auf mögliche Ursachen geschlossen werden. Woolfolk (2008) geht davon aus, dass eine Person dann ihre Selbstwirksamkeit steigert, wenn sie sich in einer Situation als kompetent erlebt. Bezogen auf den Laborbesuch spricht dies für ein zielgruppengerechtes

Kurskonzept. Offensichtlich gelingt es im Laborkurs Genetischer Fingerabdruck, Schüler weder zu über- noch zu unterfordern, was einem Kompetenzerfinden entgegenstehen würde. Wie Reinmann & Mandl (2006) betonen, entsteht Überforderung leicht durch Mangel an instruktionaler Unterstützung. Diese Unterstützung wurde im BeLL Bio durch das Versuchsskript und die Betreuer gewährleistet. Wie bereits gezeigt wurde, wird in allen Laborkursen die Versuchsanleitung als sehr transparent eingeschätzt und die Kommunikation mit dem Betreuer als relativ wichtig eingestuft. Eine Unterforderung der Schüler im Laborkurs hätte dagegen voraussichtlich eher zu Langeweile geführt (Csikszentmihalyi & Schiefele 1993). Dass dies weder im Laborkurs Genetischer Fingerabdruck, noch im Kurs LDH der Fall ist, konnte im Abschnitt 3.2.2.3.3 gezeigt werden. Dort ist dargestellt, dass Schüler nach dem Laborbesuch ein hohes aktuelles emotionales Interesse bekunden, was gegen das Empfinden von Langeweile spricht.

Nun bleibt zu diskutieren, warum zwar die Selbstwirksamkeitserwartung, Experimente durchzuführen, nicht aber die subjektiv empfundene Planungs- und Auswertungskompetenz gesteigert werden konnten. Die Laborkurse des BeLL Bio legen ihren inhaltlichen Schwerpunkt auf die Durchführung und Auswertung von Experimenten. Schüler erhalten hier die Möglichkeit, Labormethoden, welche zuvor im Rahmen des Biologieunterrichts behandelt werden, in der Praxis umzusetzen. Da die Planung des Versuchsablaufes nicht Bestandteil des Laborkurses ist, wäre eine deutliche Steigerung der entsprechenden Selbstwirksamkeitserwartung, Experimente planen zu können, überraschend gewesen. Dennoch konnte bei den männlichen Schülern im Posttest der Nebenstudie eine geringfügige, wenn auch gegenüber der Kontrollgruppe nicht signifikante, Steigerung der subjektiv empfundenen Planungskompetenz festgestellt werden. Einen Beitrag zu der etwas höheren Einschätzungen der männlichen Schüler könnte das größere Interesse männlicher Schüler an technischen Inhalten (Holstermann & Bögeholz 2007), in Verbindung mit der bereits erwähnten Tendenz, ihre Leistungsfähigkeit zu überschätzen (Helbig 2012), geliefert haben. Eventuell verleitete der vermeintlich vertraute Umgang mit den technischen Laborgeräten bei den männlichen Schülern zu der Annahme, den Versuch auch selbst leicht planen zu können. Das Argument der Überschätzung würde auch die unmittelbar im Anschluss des Laborkurses geringfügig höher eingeschätzte Auswertungskompetenz der männlichen Schüler erklären. Verwundern könnte, dass in der Subskala Auswertung nicht generell, also geschlechterunabhängig, eine deutliche Steigerung der Selbstwirksamkeit zu verzeichnen ist, so wie in der Subskala Durchführung. Eine Erklärung liefert der Zeitanteil, den die Auswertung im Laborkurs Genetischer Fingerabdruck einnimmt. Zwar ist die Auswertung der Versuchsergebnisse elementarer Bestandteil des Laborkurses, dennoch nimmt sie selbst nur einen relativ geringen Anteil der gesamten Versuchsdauer ein. Hier werden zentrale Beobachtungen an den DNA-Banden des Agarose-Gels zusammengetragen und mögliche

Fehlerquellen diskutiert. Eine weiterführende Besprechung und Deutung der Ergebnisse obliegt der jeweiligen Lehrkraft im schulischen Biologieunterricht.

Die Nachbereitung der Schülerlaborkurse im Schulunterricht erhält insbesondere in Bezug auf die Nachhaltigkeit der durch den Laborbesuch erzielten Wirkung Relevanz. Erst die Nachbereitung hilft den Schülern, Kursergebnisse zu reflektieren (Davies et al. 2008). Häufig fehlt in Schülerlaboren eine Nachbereitung der Schülerlaborkurse allerdings völlig (Brandt 2005). Folglich ist zu erklären, dass auch die durch den Laborkurs Genetischer Fingerabdruck erzielte Steigerung der empfundenen Durchführungskompetenz im Follow-up-Test nur noch geringfügig über dem Pretestniveau liegt. Dieser Rückgang wird sicherlich durch die ohnehin nur relativ geringe Frequenz von Schülerexperimenten im Biologieunterricht (Scharfenberg 2005) gefördert.

Bisher konnte die Steigerung der experimentbezogenen Selbstwirksamkeitserwartung nur für den Laborkurs Genetischer Fingerabdruck belegt werden. Inwiefern diese Steigerung global auch für andere Schülerlaborkurse beobachtet werden kann, gilt es zukünftig zu untersuchen. Weitere Studien müssen zudem zeigen, welche Variablen für die Steigerung der Selbstwirksamkeitserwartung verantwortlich sind und ob entsprechend konzipierte Laborkurse auch die wahrgenommene Planungs- und Auswertungskompetenz beeinflussen. Um die Nachhaltigkeit entsprechender Effekte sicherzustellen, bedarf es darüber hinaus der Entwicklung didaktischer Unterrichtskonzepte für den dem Laborkurs folgenden Biologieunterricht. Zu ergänzen ist, dass bisher keine Aussage darüber getroffen werden kann, ob eine derartige Steigerung der Selbstwirksamkeit nur am Lernort Schülerlabor zu erreichen ist. Denkbar ist, dass eine ähnliche Wirkung bei der Durchführung von Experimenten auch im schulischen Kontext erzielbar wäre. Zur Beantwortung dieser Fragestellung müsste ein Versuchsdesign gewählt werden, in welchem dem Laborkurs entsprechende Experimente in der Schule durchgeführt werden. Eine derartige Studie, welche den Einfluss des Lernortes auf die affektive und kognitive Dimension untersucht, wird aktuell an der Universität Bielefeld von Dreischmeier (in prep.) durchgeführt. Hier wird allerdings nicht die experimentbezogene Selbstwirksamkeit erhoben, so dass eine derartige Untersuchung noch aussteht. Allerdings ist zu berücksichtigen, dass – auch wenn entsprechende Experimente im schulischen Kontext eine ähnliche Wirkung erzielen sollten – die in Schülerlaboren durchgeführten Experimente in der Schule häufig kaum realisierbar sind. Somit wird es im schulischen Kontext in der Praxis auch kaum die Möglichkeit gegeben, zu einer realistischen Kompetenzeinschätzung bezüglich der Durchführung moderner, forschungsnaher Versuche beizutragen. In Hinblick darauf, dass der Selbsteinschätzung ein positiver Einfluss auf Verhalten, Leistungsmotivation und Interesse (Bandura 1982, Eccles et al. 1983, Köller et al. 2006) zugesprochen wird, erscheint das Potential von Schülerlaboren somit enorm.

Wurde bisher nur die durch den Laborbesuch bedingte quantitative Steigerung der Selbstwirksamkeit als bedeutsam bewertet, so deuten sich in den bereits vorgestellten Korrelationen der postFSWEx.V1-Daten mit den Messwerten aus Pre- und Follow-up-Test auch qualitative Veränderung in der Selbstwirksamkeitserwartung an. Hier zeigt sich, dass die höchsten Korrelationen in der Regel zwischen Pre- und Follow-up-Test-Ergebnissen vorliegen, während die unmittelbar nach dem Laborbesuch erhobenen Posttest-Daten zum Teil deutlich geringer mit ihnen korrelieren. Schüler schätzen ihre Selbstwirksamkeit im Laborkurs also nicht unbedingt so ein, wie kurz vor oder Monate nach dem Laborbesuch. Diese Beobachtung gilt besonders ausgeprägt für die Subskala Durchführung, mit Korrelationen $r < 0,2$ für die Pre- und Posttestdaten. Auffällig ist für diese Subskala, dass die Korrelationen der Posttestdaten zu den Follow-up-Testdaten höher ausfallen. Insbesondere im Kurs LDH fallen diese mit $r = 0,498$ sogar höher aus, als die Pre-Follow-up-Korrelationen, was darauf hindeutet, dass Schülerinnen und Schüler ihre am Laborkurstag erworbene Einschätzung auch nachhaltig übernehmen. Ähnliches lässt sich auch in der Subskala Auswertung im LDH-Kurs beobachten. Auch hier sind eine moderate Pre-Posttest-Korrelation und eine deutlich höhere Post-Follow-up-Test-Korrelation festzustellen. Lediglich für die Subskala Planung lässt sich kein nennenswerter Unterschied zwischen Pre-Post- und Post-Follow-up-Korrelation feststellen. In dieser Subskala führt der Laborbesuch somit zu keiner qualitativen Veränderung.

Qualitativ meint hier also nicht eine höhere, sondern – möglicherweise – eine realistischere Selbsteinschätzung. Wie Engeln (2004) feststellt, haben Schüler wenig Erfahrung im selbständigen Experimentieren. Entsprechend schwierig sollte ihnen eine realistische Einschätzung ihrer Experimentierkompetenz fallen. Da der Schwerpunkt in allen BeLL Bio-Kursen auf dem praktischen Durchführen von Experimenten liegt, deutet sich in dieser Subskala auch die deutlichste qualitative Änderung an. Im Laborkurs Genetischer Fingerabdruck nimmt die Auswertung des Agarosegels nur relativ wenig Zeit in Anspruch. Deutlich größer ist der Auswertungsanteil im Laborkurs LDH. Hier gilt es für die Schüler, Diagramme richtig zu lesen und Konzentrationen zu berechnen. Entsprechend verwundert es nicht, dass im Laborkurs Genetischer Fingerabdruck die Pre-Post-Korrelation ähnlich wie die Post-Follow-up-Korrelation ausfällt. Im LDH Kurs dagegen fällt die Post-Follow-up-Korrelation deutlich höher aus als die Pre-Posttest-Korrelation, was auf eine nachhaltig geänderte Einschätzung der Auswertungskompetenz hindeutet. Dem Beitrag eines Schülerlaborkurses zu einer realistischeren Selbsteinschätzung kann sicherlich eine mindestens ebenso wichtige Bedeutung beigemessen werden, wie die im vorherigen Abschnitt diskutierte quantitative Steigerung.

4.2.2.2 Fähigkeitsselbstkonzept und Interesse

Schülerlaborbesuchen wird das Potential zugesprochen, gleichermaßen das Selbstkonzept von Jungen und Mädchen zu fördern (Euler 2009a/b). Weiterhin stellt Engeln (2004) fest, dass verschiedene Schülerlabore sich bedeutend in ihrem Potential unterscheiden, aktuelles Interesse zu bewirken. Köller et al. (2006) stellen positive Effekte des schulischen Selbstkonzeptes und des Interesses auf den Wissenszuwachs fest. Entsprechend setzt sich auch das BeLL Bio zum Ziel, Interesse für die Naturwissenschaft Biologie zu wecken und Selbstvertrauen im Umgang mit biologischen Sachverhalten zu bewirken. Daraus ergibt sich die Fragestellung:

Welchen Einfluss hat ein einmaliger Laborbesuch im Lernort BeLL Bio auf das aktuelle Interesse und das biologische Fähigkeitsselbstkonzept?

Das Fähigkeitsselbstkonzept stellte sich in der vorliegenden Untersuchung als nicht statisch heraus. Insbesondere in der Versuchsgruppe Genetischer Fingerabdruck konnte eine nachweisliche Erhöhung des Fähigkeitsselbstkonzeptes durch den Laborbesuch belegt werden. Zwar ist diese nur kurzfristig und nur in diesem Kurs statistisch belegbar, doch deuten auch die Beobachtungen in den anderen beiden BeLL Bio-Kursen augenscheinlich an, dass bereits ein einmaliger Besuch im Schülerlabor zu einer geringfügigen, kurzfristigen Änderung des Fähigkeitsselbstkonzeptes beitragen kann. Eine bedeutsame Abhängigkeit vom Geschlecht oder Migrationshintergrund ist in keinem der Kurse zu beobachten. Damit stützt die vorliegende Studie die Arbeiten von Brandt (2005) und Pawek (2009), welche ebenfalls nach einem einmaligen Laborbesuch eine kurzfristige, geschlechterunabhängige Steigerung des fachlichen Selbstkonzeptes bei Schülern beobachten. Darüber hinaus kann im Kurs Genetischer Fingerabdruck eine nachhaltigere Steigerung des Fähigkeitsselbstkonzeptes für Schüler mit naturwissenschaftlicher beruflicher Orientierung im Vergleich zu Schülern ohne derartige Ausrichtung belegt werden. Diese scheint allerdings kursabhängig zu sein, da im LDH-Kurs für diese Schülergruppe ebenfalls – genau wie für die Schüler mit nicht naturwissenschaftlicher Berufsorientierung – nur eine kurzfristige Steigerung zu beobachten ist. Ein Grund für die nachhaltigere Erhöhung des Fähigkeitsselbstkonzeptes im Laborkurs Genetischer Fingerabdruck bei den naturwissenschaftlich Orientierten könnte ihr etwas höheres epistemisches Interesse (vgl. Punkt 3.2.2.3.3) im Anschluss an den Laborkurs sein. Die naturwissenschaftlich orientierten Schüler geben somit deutlicher an, sich auch in ihrer Freizeit noch mit den Laborkursinhalten auseinander setzen zu wollen. In Verbindung mit ihrem höheren dispositionalen Fachinteresse könnte dies zu einer intensiveren Auseinandersetzung mit den Laborkursthemen über den Laborbesuch hinaus führen, was auch im Biologieunterricht zu entsprechenden Erfolgserlebnissen und damit zu einem nachhaltig höheren Selbstkonzept verhelfen würde. Interessante Beobachtungen ergaben sich auch im Vergleich der Schülertypen. Die Praktiker erfahren augenscheinlich sowohl kurz, als auch langfristig

eine etwas höhere Steigerung ihres Fähigkeitsselbstkonzeptes, als die Theoretiker und insbesondere als die Allrounder. Allerdings ließ sich statistisch nur die nachhaltige Steigerung des Fähigkeitsselbstkonzeptes der Praktiker gegenüber den Allroundern im LDH-Kurs belegen. Denkbare Ursache dafür ist, dass den Praktikern im Schulunterricht gewöhnlich kaum Kompetenzerlebnisse ermöglicht werden, da hier wenig selbst experimentiert wird (Seidel et al. 2007). Fachlich befinden sie sich in einem deutlich stärkeren sozialen Umfeld und schätzen daher ihr eigenes Selbstkonzept relativ gering ein (*Big-Fish-Little-Pond-Effekt*, Marsh 2005). Umso bedeutsamer muss es für die Praktiker sein, im Schülerlabor entsprechend ihres Interesses am Experimentieren und ihrer relativ hoch eingeschätzten Durchführungskompetenz selbst experimentieren zu dürfen.

In Bezug auf das aktuelle Interesse konnte festgestellt werden, dass Schüler im Anschluss an die BeLL Bio-Besuche hohes emotionales Interesse bekunden. Dieses fällt allerdings nicht höher aus, als das emotionale Interesse, welches von der Kontrollgruppe bezüglich des Experimentierens im schulischen Biologieunterricht geäußert wird. Das wertebezogene Interesse ist ähnlich hoch, wie das emotionale Interesse, während das epistemische Interesse sowohl in schulischem Biologieunterricht, als auch im Schülerlabor deutlich geringer ausfällt. Auch Engeln (2004), Guderian (2007) und Pawek (2009) stellen das geringste Interesse von Schülern nach einem Schülerlaborbesuch in der epistemischen Komponente fest. Die Bereitschaft, sich über den Schülerlaborbesuch hinaus in der Freizeit mit den Laborkursthemen zu beschäftigen, ist also generell bei Schülern mäßig stark ausgeprägt. Unter Berücksichtigung, dass im Lauf der Sekundarstufe ein genereller Interessenrückgang bei Schülern in der Biologie zu verzeichnen ist (Löwe 1987), verwundert dies nicht. Während sich die drei Interessenkomponenten in der Kontrollgruppe als zeitlich stabil erwies, konnte in den Versuchsgruppen Genetischer Fingerabdruck und LDH ein deutliches Absinken des wertebezogenen und epistemischen Interesses zum Follow-up-Test hin beobachtet werden. Das emotionale Interesse stellte sich dagegen auch bei den Laborbesuchern als langfristig stabil heraus. Ähnliche Beobachtungen macht auch Engeln (2004). Sie stellt innerhalb von zwölf Wochen nach einem Laborbesuch einen geringfügigen Rückgang des epistemischen, aber auch des emotionalen Interesses fest. Als mögliche Gründe für den Interessenrückgang führt sie an, dass einige Wochen nach dem Laborbesuch bereits andere Eindrücke die Wirkung des Labores überdecken. Als eine weitere mögliche Erklärung für den Rückgang des epistemischen Interesses nennt sie die Itemformulierung dieser Subskala. Während im ersten Posttest alle Items als Intentionen formuliert sind, wurde im Follow-up-Test nicht nach der Absicht, sondern nach manifestiertem Verhalten gefragt (Engeln 2004).

Bedeutsame Unterschiede in der Entwicklung des aktuellen Interesses zwischen männlichen und weiblichen Schülern sind nicht festzustellen. Für Schüler mit Migrationshintergrund ist im Laborkurs

Genetischer Fingerabdruck ein etwas höherer Rückgang des emotionalen Interesses als bei den Schülern ohne Migrationshintergrund zu beobachten. Eine deutliche Abhängigkeit vom Migrationshintergrund hätte verwundert, da in PISA 2006 zwischen Jugendlichen mit und ohne Migrationshintergrund keine Unterschiede in der Ausprägung ihres naturwissenschaftlichen Interesses beobachtet werden konnten (Prenzel 2007). Deutlich höher fällt in der vorliegenden Studie dagegen das aktuelle Interesse der Schüler mit naturwissenschaftlicher Berufsorientierung aus. Sie weisen sowohl unmittelbar nach dem Laborbesuch, als auch noch nach über zwei Monaten ein deutlich höheres Interesse auf, als die nicht naturwissenschaftlich orientierten Schüler. Ursächlich dafür könnte das ohnehin bei den naturwissenschaftlich Orientierten bedeutend höhere dispositionale Fachinteresse am Unterrichtsfach Biologie sein. Zudem belegt das höhere wertebezogene Interesse dieser Schüler, dass sie den Inhalten der Laborkurse in Hinblick auf ihre Berufspläne offensichtlich einen hohen Nutzen zusprechen. Diese Beobachtung erklärt auch den Befund, dass naturwissenschaftlich orientierte Schüler deutlich stärker zustimmen, wenn sie gefragt werden, ob der Laborbesuch ihr Interesse für naturwissenschaftliche Berufe gesteigert hat.

Ähnlich zeigt sich auch bei der Differenzierung der Schülertypen, dass die fachlich sehr interessierten Allrounder in allen Komponenten das entsprechend höchste aktuelle Interesse bekunden. Im LDH-Kurs folgen die Theoretiker mit einer etwas geringeren Interessenausprägung. Das niedrigste aktuelle Interesse wird in diesem Laborkurs von den Praktikern bekundet. Ein ähnliches Bild ergibt sich für den Laborkurs Genetischer Fingerabdruck für die emotionale Interessenskomponente. In der epistemischen und wertebezogenen Komponente werden hier allerdings die Theoretiker von den Praktikern im Grad ihres aktuellen Interesses übertroffen. Eine mögliche Erklärung der Unterschiede zwischen beiden Laborkursen ist im Kursinhalt zu suchen. Im Laborkurs LDH führen die Schüler photometrische Messungen durch. Sie dokumentieren ihre Ergebnisse, werten diese in einer Tabellenkalkulation aus, erstellen Grafiken und führen Berechnungen durch. Dieser hohe Anteil an Kopfarbeit spricht vermutlich eher die fachlich Interessierten Allrounder und Theoretiker an. Im Laborkurs Genetischer Fingerabdruck ist dieser theoretische Anteil geringer. Hier arbeiten die Schüler mit Mikroliterpipette, Zentrifugen, PCR-Geräten und Gelkammern. Der theoretischere Anteil beschränkt sich hier auf die Auswertung des Agarose-Gels am Ende des Laborkurses.

Zu diskutieren bleibt, warum sich das Interesse der Kontrollgruppe über die Messzeitpunkte hin nicht ändert. Generell hat auf das aktuelle Interesse die Interessantheit der Situation, aber auch das dispositionale Interesse Einfluss. Da die Kontrollgruppe „unbehandelt“ blieb, also im Untersuchungszeitraum kein Schülerlabor besuchte, konnte bei ihr nicht das situationsabhängige aktuelle Interesse, sondern vielmehr das dispositionale Interesse erhoben werden. Dieses gilt als relativ stabiles Persönlichkeitsmerkmal (vgl. Vogt 2007), so dass hier keine Änderung zu erwarten

gewesen wäre. Das dispositionale Interesse der Kontrollgruppe dient hier als Baseline zur Einschätzung der Größenordnung des Interessenrückgangs in den Versuchsgruppen. Die Unveränderlichkeit des Interesses in der Kontrollgruppe verbunden mit der zeitlichen Veränderung des Interesses in den Versuchsgruppen weist darauf hin, dass in den Versuchsgruppen tatsächlich das aktuelle Interesse erfasst wurde. Ob eine nachhaltigere Förderung des aktuellen Interesses allein durch mehrmaligen Laborbesuch zu erreichen ist, ist fraglich. Guderian (2007) untersucht die Änderung der drei Komponenten des aktuellen Interesses über drei Schülerlaborbesuche hinweg. Die Besuche folgen im Abstand von fünf bis sechs Wochen aufeinander. Bei Fünftklässlern stellt er nach jedem Laborbesuch in allen Interessenkomponenten einen kurzfristigen Anstieg und einige Wochen später wieder einen deutlichen Rückgang fest. Eine längerfristige Wirkung im Sinne einer nachhaltigen Interessensteigerung kann nur erreicht werden, wenn es gelingt, dass die Schüler sich mit dem Lerngegenstand identifizieren (Krapp 2010). Dies scheint im Rahmen einfacher oder mehrfacher, in der Regel nur kurzzeitiger Schülerlaborbesuche kaum möglich zu sein. Guderian et al. (2006a) folgern daher, dass es im Anschluss an den Besuch eines Schülerlabors in den Händen der Lehrkraft liegt, aus dem aktuellen Interesse nachhaltiges Interesse aufzubauen.

4.2.2.3 Wissenserwerb

Das BeLL Bio möchte explizit nicht nur eine affektive Wirkung, sondern auch einen Lerneffekt bei den Schülern bewirken. Ziel ist es, den Schülern ein nachhaltiges Lernerlebnis zu ermöglichen.

Es gilt daher zu prüfen, ob die im BeLL Bio angebotenen Laborkurse zu einem nachhaltigen kognitiven Wissenserwerb beitragen.

Für die Entwicklung des Fachwissens über die Messzeitpunkte kann festgehalten werden, dass es allen drei Laborkursen gelingt, einen höchst signifikanten Wissenszuwachs bei den Kursteilnehmerinnen und Kursteilnehmern zu bewirken. Dieser kann für die Laborkurse Genetischer Fingerabdruck und LDH als nachhaltig belegt werden. Die differenzierteren Betrachtungen zeigen, dass weibliche und männliche Schüler, genau wie Schüler mit und ohne Migrationshintergrund sowie naturwissenschaftlich und nicht naturwissenschaftlich orientierte gleichermaßen gefördert werden. Dies ist insbesondere hinsichtlich der beruflichen Orientierung bemerkenswert, da im Laborkurs Genetischer Fingerabdruck ein deutlich höheres Vorwissen der naturwissenschaftlich orientierten Schüler festgestellt wurde. Beim Vergleich der drei sehr unterschiedlichen Schülertypen wird nur im Falle des Kurses Genetischer Fingerabdruck ein Unterschied im Vorwissen zugunsten der Theoretiker und Allrounder erkennbar. Sowohl im Kurs Genetischer Fingerabdruck, als auch im LDH-Kurs deutet sich ein etwas höherer Wissensgewinn der Allrounder und Theoretiker gegenüber den Praktikern an, was statistisch allerdings nur schwach gestützt wird. Im Einzeller-Kurs sind keine Unterschiede im

Lernzuwachs der drei Gruppen zu beobachten. Die drei sehr heterogenen Schülertypen werden somit im BeLL Bio relativ homogen gefördert.

Auch Scharfenberg (2005) stellt eine relativ gleichmäßige kognitive Förderung von Schülern im Schülerlabor fest. Er untersuchte den Einfluss des Lernortes auf den kognitiven Lernzuwachs, indem er unter anderem den durch einen Laborbesuch bedingten Lernzuwachs einer nicht experimentierenden und einer experimentierenden Laborgruppe vergleicht. Dabei stellt er fest, dass sich bereits die Lernumgebung Labor alleine, ohne dass hier Experimente durchgeführt werden müssen, förderlich auf den Lernzuwachs auswirkt. Weiterhin stellt er fest, dass die Lernleistung von Schülern am Lernort Labor deutlich schwächer mit der Schulnote korreliert, als dies bei der Lernleistung der schulischen Vergleichsgruppe der Fall ist. Dies deutet darauf hin, dass es Schülerlaboren im Vergleich zum schulischen Unterricht besser gelingt, leistungsschwächere Schüler zu fördern. Auch Engeln und Euler (2004) sehen in Schülerlaboren das Potential „[...] Problemgruppen [...]“ (Engeln & Euler 2004) unter den Schülern zu erreichen. Gleichzeitig merkt Scharfenberg (2005) allerdings auch an, dass es sein könnte, dass leistungsstärkere Schüler unter Umständen unter ihren Möglichkeiten bleiben. Da in der Praxis im Schülerlabor wohl kaum nur theoretisch gearbeitet wird, erschiene der Vergleich einer im Labor und in der Schule experimentierenden Schülergruppe von deutlicherem Interesse, um den Einfluss des Lernortes auf den Lernerfolg ermitteln zu können. Eine solche Studie wird aktuell im Bielefelder Schülerlabor teutolab-Biotechnologie von Dreischmeier (in prep.) durchgeführt. Zwar wird in der hier vorliegenden Arbeit kein derartiger Vergleich vorgenommen, dennoch deutet sich auch hier eine weitgehend von Geschlecht, Migrationshintergrund, beruflicher Orientierung und Schülertypus unabhängige kognitive Förderung an. Dass diese noch Wochen nach dem Laborbesuch nachhaltig zu belegen ist, ist sicherlich nicht allein durch nachbereitenden Unterricht in der Schule bedingt. In der Regel wird der BeLL Bio-Besuch nämlich erst zum Ende einer Unterrichtseinheit in den Unterricht integriert, so dass im Anschluss an den Laborbesuch im Biologieunterricht häufig zeitnah ein neues Thema begonnen wird. John und Thinesse-Demel (2004) gehen davon aus, dass Lernen in einer solchen Umgebung von Nachhaltigkeit geprägt ist, die Voraussetzung für bessere Bedingungen im Leben und im Beruf schafft. Entsprechende Bedeutung sollte der Wahrnehmung der Lernumgebung Labor durch die Schüler zukommen. Diese wird im folgenden Punkt diskutiert.

4.2.2.4 Wahrnehmung der Lernumgebung Labor

Im Rahmen der PISA 2009-Erhebung wurde festgestellt, dass die Leistungsspannen im Bereich der naturwissenschaftlichen Kompetenz zwischen sehr starken und sehr schwachen Schülern in Deutschland relativ groß ausfällt (Rönnebeck et al. 2010). Umso wichtiger erscheint die Beobachtung, dass die Wirkung einmaliger Laborbesuche im BeLL Bio – wie in den letzten Abschnitten diskutiert

wurde – sowohl auf affektiver, als auch kognitiver Ebene weitgehend unabhängig von Geschlecht, Migrationshintergrund, naturwissenschaftlicher Orientierung und Schülertypus ausfällt. Ob sich diese Homogenität auch in der Wahrnehmung des Lernortes Schülerlabor widerspiegelt, soll unter Beantwortung der folgenden Fragestellung geklärt werden.

Gelingt es den BeLL Bio-Kursen, ein heterogenes Schülerfeld unabhängig von Geschlecht, Migrationshintergrund und beruflicher Orientierung gleichermaßen anzusprechen?

Zunächst wird hier die Wahrnehmung der Laborumgebung in den Versuchsgruppen im Vergleich zur Wahrnehmung des Lernortes Biologieraum durch die Kontrollgruppe erörtert. Im Anschluss werden Unterschiede zwischen den Subgruppen diskutiert.

Hier kann festgehalten werden, dass sich die drei Laborkurse des BeLL Bio durch eine als deutlich höher empfundene Authentizität, eine höhere Einschätzung der Bedeutung des Betreuers, sowie eine positivere Atmosphäre gegenüber dem schulischen Biologieunterricht auszeichnen. Diese Beobachtungen fallen entsprechend der Erwartungen aus, gründet sich das Konzept des BeLL Bio doch auf einem gemäßigt konstruktivistischen Ansatz (vgl. 1.1.2.1.4). Dieser sieht die Authentizität der Lernumgebung als Notwendigkeit für erfolgreiches Lernen an. Im BeLL Bio wird diese durch die Arbeit in einem forschungsnahen Laborraum der Bergischen Universität Wuppertal, den Umgang mit echten Laborgeräten und Problemstellungen, sowie den originalen Kontakt mit Studierenden und Wissenschaftlern gewährleistet. Wie die hohen Einschätzungen der Schüler zeigen, trägt die Lernumgebung Laborraum aber nicht nur zur Authentizität bei, sie erzeugt auch gemeinsam mit den sozialen Interaktionen im Labor (vgl. Pfister 2011) eine angenehme Lernatmosphäre. Dies gelingt ihm offensichtlich besser, als dem Biologieraum in der Schule. Dazu passt die Beobachtung, dass die Schüler den Gesprächen mit dem Kursbetreuer im Labor mehr Bedeutung beimessen, als die Schüler der Kontrollgruppe dies für die Gespräche mit ihrem Biologielehrer tun. Hier bestätigt sich die Annahme von Anton und Hock (2003), dass Schülerlabore eine Chance darstellen, die Berührungängste der Schüler mit Wissenschaft zu verringern. Sicherlich trägt dazu auch die geringe Altersdistanz der Oberstufenschüler zu den studentischen Hilfskräften im BeLL Bio bei. Auch die als recht gut empfundene Zusammenarbeit mit den Mitschülerinnen und Mitschülern in den drei Laborkursen sprechen für eine gute Lernatmosphäre. Die Qualität der Zusammenarbeit wird von den Schülern der Kontrollgruppe im Biologieunterricht allerdings ähnlich hoch empfunden.

Interessant ist auch die Beobachtung, dass die Schüler der Kontrollgruppe im regulären Biologieunterricht eine höhere Offenheit bei gleichzeitig höherem Druck empfinden, als die Schüler der Versuchsgruppen LDH und Genetischer Fingerabdruck. Ähnlich hohe Offenheit und Druck wie die Kontrollgruppe nehmen auch die Schüler der Versuchsgruppe Einzeller wahr. Innerhalb der

Versuchsgruppen kann zudem ein moderater positiver Zusammenhang von Offenheit und Druck festgestellt werden. Generell ist der positive Zusammenhang von Offenheit und Druck verständlich, denn gerade im Zusammenhang mit offen gehaltenen Experimentalaufgaben ist anzunehmen, dass sich Schüler aufgrund ihrer mangelnden Experimentiererfahrung schnell überfordert fühlen (Kirchner & Priemer 2007). Der im Laborkurs Einzeller als höher empfundene Druck könnte daraus resultieren, dass sich hier das methodische Konzept dieses Laborkurses von den Kursen LDH und Genetischer Fingerabdruck unterscheidet. Während die Schüler in den zuletzt genannten Kursen einen einzigen Versuch in relativ festgelegter Reihenfolge durchführen, arbeiten die Schüler im Kurs Einzeller zum Teil an Stationen in nicht vorgegebener Reihenfolge. Engeln (2004) stellt fest, dass Offenheit, Verständlichkeit und Authentizität in einem Spannungsverhältnis stehen. Sie empfiehlt daher, dass Schülerlabore sich um Authentizität und Verständlichkeit bemühen sollen, auch wenn dies auf Kosten der Offenheit geht. Eine hohe Verständlichkeit belegt in den drei Versuchsgruppen die als sehr hoch eingeschätzte Transparenz der Versuchsanleitung.

Ein deutlicher Unterschied in der Laborwahrnehmung findet sich zwischen den Versuchsgruppen lediglich im wahrgenommenen Alltagsbezug. Hier geben die Schüler im Laborkurs Genetischer Fingerabdruck nachweisbar höhere Werte an, als in den anderen beiden Laborkursen und der Kontrollgruppe. Offensichtlich gelingt es diesem Laborkurs am deutlichsten an die Lebenswelt der Schülerinnen und Schüler anzuknüpfen. Grund hierfür dürfte das in den letzten Jahren stark zunehmende Interesse von Jugendlichen am Thema Rechtsmedizin sein. Ausgelöst wurde dieses mit dem Aufkommen einschlägiger TV-Serien wie *CSI-Den Tätern auf der Spur*. Keuneke et al. (2010) fassen diese Entwicklung auch als „CSI-Effekt“ zusammen. Sie stellen im Rahmen einer qualitativen Studie fest, dass Jugendliche den Wunsch verspüren, das medial vermittelte verzerrte Bild der Rechtsmedizin mit der Realität abzugleichen. Hier knüpft der Laborkurs Genetischer Fingerabdruck thematisch unmittelbar an, was zu einem entsprechend hoch eingeschätzten Alltagsbezug führt. Der Laborkurs LDH führt zwar über das Thema Sport an das Thema Stoffwechsel heran, offensichtlich ist hier der Bezug zur Lebenswelt der Schüler aber nicht derart offensichtlich. Da die sportliche Aktivität und die dem Thema Sport beigemessene Relevanz von Schüler zu Schüler deutlich variieren wird, ist diese Beobachtung nachvollziehbar. Etwas geringer als im LDH-Kurs fällt der wahrgenommene Alltagsbezug im Laborkurs Einzeller aus. Ein Grund hierfür könnte sein, dass sich in den thematischen Stationen nur singuläre Anknüpfungen an den Alltag der Schüler finden.

Wie die Subgruppenvergleiche zeigen, wird der Lernort BeLL Bio von Schülerinnen und Schülern unabhängig von ihrem Geschlecht geschätzt. Diese Beobachtungen decken sich mit Pawek (2009). Auch er stellt nur geringfügige Unterschiede zwischen den Geschlechtern in der Wahrnehmung der Laborvariablen fest. Sowohl im Laborkurs LDH, als auch im Kurs Genetischer Fingerabdruck konnte

darüber hinaus beobachtet werden, dass Schüler mit Migrationshintergrund eine höhere Offenheit, im Kurs Genetischer Fingerabdruck verbunden mit höherem Druck, empfanden. Denkbar wären auf den Migrationshintergrund zurückzuführende sprachliche Probleme. Ob dies tatsächlich der Grund ist, könnte gegebenenfalls eine Aufschlüsselung nach der Generation des Migrationshintergrundes klären (Prenzel 2007). Da in der Versuchsgruppe Genetischer Fingerabdruck allerdings nur neun und im Kurs LDH nur drei Schüler einen Migrationshintergrund erster Generation aufweisen, also selbst außerhalb von Deutschland geboren wurden, scheint eine solche Aufschlüsselung nicht sinnvoll. Größere Verständnisprobleme der Schüler mit Migrationshintergrund scheint es allerdings nicht zu geben, da die Versuchsanleitung von diesen Schülern ebenfalls als transparent bewertet wurde.

Während sich die naturwissenschaftlich orientierten Schüler in ihrer persönlichen Disposition deutlich von den nicht naturwissenschaftlich ausgerichteten Schülern unterscheiden, spielt die berufliche Orientierung für die Wahrnehmung der Lernumgebung im Laborkurs Genetischer Fingerabdruck keine und im Laborkurs LDH nur eine mäßige Rolle. In der schulischen Kontrollgruppe sind dagegen deutlichere Unterschiede zwischen naturwissenschaftlich und nicht naturwissenschaftlich orientierten Schülern zu beobachten. In der Kontrollgruppe erkennen diejenigen mit naturwissenschaftlichem Berufs- oder Studienziel für den schulischen Biologieunterricht einen deutlich höheren Alltagsbezug und empfinden die Atmosphäre im Biologieraum als deutlich angenehmer, als die nicht naturwissenschaftlich orientierten Schüler. Ähnliche Beobachtungen ergeben sich im Vergleich der drei Schülertypen. Im Laborkurs Einzeller und im Kurs Genetischer Fingerabdruck ist die Wahrnehmung des Laborkurses kaum abhängig vom Schülertypus. Alle Schülertypen schätzen in diesen Kursen die Laborvariablen sehr ähnlich ein. Deutlichere Unterschiede sind im Laborkurs LDH zu beobachten. Hier fällt insbesondere der von den Allroundern gegenüber den Praktikern und Theoretikern höher empfundene Alltagsbezug auf. Sehr deutliche Unterschiede in der Wahrnehmung der Lernumgebung sind in der Kontrollgruppe für die Variablen Druck, Zusammenarbeit, Atmosphäre und Betreuer festzustellen. Die Praktiker fallen in der Kontrollgruppe insbesondere dadurch auf, dass sie der Kommunikation mit ihrem Biologielehrer nur wenig Bedeutung beimessen und die Atmosphäre im Biologieraum weniger positiv beurteilen, als die anderen beiden Schülertypen. In den Laborkursen ist ein solcher Unterschied dagegen nicht zu beobachten. Der Schülertypus scheint somit insbesondere im schulischen Biologieunterricht, aber auch im Laborkurs LDH nicht unbedeutend für die Wahrnehmung der Lernumgebung zu sein. In den Laborkursen Genetischer Fingerabdruck und Einzeller spielt der Schülertypus dagegen keine bedeutende Rolle. Ein Grund für die homogenere Förderung verschiedener Schülertypen am Lernort Schülerlabor könnte die in der Regel für alle Schüler neue Lernumgebung sein. Die Andersartigkeit zum gewohnten Schulumfeld kann dazu beitragen, dass die Schüler sich öffnen und das gewohnte Verhaltensrepertoire ablegen (Hoffmeister 2003). So ist auch zu erklären, dass die fachlich weniger

interessierten Praktiker im Schülerlabor eher den Kontakt zu betreuenden Personen suchen, als dies im Biologieunterricht der Fall ist. Der gegenüber den Allroundern deutlich niedriger empfundene Alltagsbezug der Praktiker und Theoretiker im LDH-Kurs könnte darauf zurückzuführen sein, dass in diesem Kurs der Alltagsbezug nicht so offensichtlich ist, wie im Kurs Genetischer Fingerabdruck. Dieser wird zwar über die Anknüpfung an die sportliche Leistungsdiagnostik hergestellt, die Anknüpfung des Themas Enzymatik und Stoffwechsel könnte aber Erfahrungen aus Alltag und Schule voraussetzen, die nur bei den fachlich besonderes interessierten Allroundern zu finden sind.

Resümierend kann somit festgehalten werden, dass es insbesondere in den Laborkursen Genetischer Fingerabdruck und Einzeller sehr gut gelingt, auch eine sehr heterogene Schülerschaft zu fördern. Damit heben sich diese Kurse vom schulischen Unterricht deutlich ab. Eine Optimierung des Alltagsbezuges im Kurs LDH wäre zu überdenken. Zu berücksichtigen ist, dass der Vergleich der Kontrollgruppe mit den Versuchsgruppen im Labor nur mit Vorsicht zu interpretieren ist, denn die „unbehandelte“ Kontrollgruppe wurde bezüglich ihrer Einstellung zum Biologieraum losgelöst von einer konkreten Experimentiersituation befragt. Weitere vergleichende Studien der Lernorte Schule und Schülerlabor – wie die aktuell von Dreischmeier (in prep.) durchgeführte – sind notwendig, um diese Beobachtungen abzusichern. Dennoch erlaubt der hier vorgenommene Vergleich eine erste Einschätzung der Laborwahrnehmung in Relation zur Wahrnehmung des schulischen Biologieunterrichts.

4.2.2.5 Modell zur affektiven Wirkung des BeLL Bio-Besuches

Die bisherige Diskussion beschränkte sich auf die singuläre Betrachtung einzelner Wirkungen und Einflussgrößen am Lernort Schülerlabor. Das im Rahmen der Fragestellung (vgl. Punkt 1.2.5) in Anlehnung an Pawek (2009) entwickelte Modell zur affektiven Wirkung des BeLL Bio-Besuches, stellt viele der beschriebenen Größen in kausale Zusammenhänge. Das aus der Theorie hergeleitete Modell führte zu folgender Fragestellung:

Wie gut repräsentiert das aufgestellte Hypothesennetz die im BeLL Bio erhobenen Daten affektiver Einflussgrößen?

Eine ausführliche Prüfung und Erklärung des Modells wurde bereits unter Punkt 1.2.5 vorgenommen, so dass hier nur wenige Aspekte ergänzt werden sollen. Das Modell erwies sich als plausibel, wenn auch einige der theoretisch hergeleiteten Kausalbeziehungen anhand der vorliegenden Daten nicht bestätigt werden konnten. Insbesondere die Einflüsse der im Modell aufgeführten dispositionalen Variablen auf die Wahrnehmung des Laborbesuches fielen nur gering aus. Lediglich ein moderater Einfluss des dispositionalen Fachwissens auf die Laborwahrnehmung konnte festgestellt werden.

Dieser Befund unterstreicht die Feststellungen aus dem vorherigen Abschnitt: Offensichtlich gelingt es dem BeLL Bio, Schüler weitgehend unabhängig von ihrer affektiven Veranlagung anzusprechen. Umso wichtiger erscheint die gelungene Konzeption der im Schülerlabor angebotenen Kurskonzepte. Ähnliche Beobachtungen macht auch Pawek (2009). Er stellt ebenfalls einen mäßigen Einfluss des dispositionalen Interesses auf den Laborbesuch, aber keinen Einfluss des dispositionalen Fähigkeitsselbstkonzeptes auf die Laborwahrnehmung fest.

Im Modell der vorliegenden Studie bestätigt sich ein signifikanter Einfluss des Laborbesuches auf das aktuelle Interesse, sowie die unmittelbar nach dem Laborbesuch empfundene experimentbezogene Selbstwirksamkeitserwartung. Der Einfluss des Laborbesuches auf das aktuelle Interesse erscheint umso bedeutsamer, da das dispositionale Fach- und Sachinteresse nur einen sehr geringen Einfluss ausübt. Sehr stark hängt dagegen das biologische Fähigkeitsselbstkonzept im Anschluss an den Laborbesuch vom dispositionalen Selbstkonzept ab. Auch diese Beobachtungen ähneln denen von Pawek (2009). Dieser stellte anhand seines Modells über die Interessenförderung durch Schülerlabore fest, dass das dispositionale Fähigkeitsselbstkonzept wesentlich das Fähigkeitsselbstkonzept nach der Laborveranstaltung beeinflusst, während die Laborveranstaltung selbst kaum einen Einfluss darauf ausübt. Weiterhin beobachtet er, dass der Einfluss des Laborbesuches auf das aktuelle Interesse so bedeutsam ist, wie der Einfluss des dispositionalen Interesses (Pawek 2009).

Der Einfluss des Laborbesuches auf die experimentbezogene Selbstwirksamkeitserwartung nach dem Laborbesuch ist in der vorliegenden Studie ähnlich stark, wie die Bedeutung der persönlichen Disposition. Auch wenn Selbstkonzept und Selbstwirksamkeit offensichtlich nicht statisch sind, ist der Einfluss des Laborbesuches auf diese Variablen also geringer. Wie die Ergebnisse der Nebenstudie zeigen, gilt dies allerdings nur für die Subskalen Planung und Auswertung. Ein Einfluss auf die empfundene Durchführungskompetenz ist – zumindest im Laborkurs Genetischer Fingerabdruck – durchaus gegeben. Auch im Modell werden allerdings nur geringe nachhaltige Effekte belegt. Interessant wäre zukünftig eine Anwendung des Modells auf Daten, welche unter systematischer Nachbereitung des Laborunterrichtes generiert wurden. Wie bereits diskutiert wurde, fehlt diese häufig im Unterricht (Engeln et al. 2003).

Entsprechend gering fällt auch die durch den Laborbesuch bedingte Einstellungsänderung für naturwissenschaftliche Berufe aus. Ein indirekter Effekt ist allerdings über das im Follow-up-Test noch relativ hohe aktuelle Interesse festzustellen. Deutlichen Einfluss auf diese Beobachtung könnte der Erhebungszeitpunkt der Variable Berufsinteresse haben. Das entsprechende Messinstrument wurde erst im Follow-up-Test eingesetzt und erfragt nachträglich, ob eine durch den Laborbesuch bedingte Interessensteigerung vorliegt. Wie das im Anschluss zum Follow-up-Test hin sinkende,

wertebezogene und epistemische Interesse zeigt, ist anzunehmen, dass eine Abfrage der Wirkung des Laborbesuches unmittelbar nach dem Labor zu höheren Einschätzungen der Schüler geführt hätte. Umso mehr ist es als beachtlich anzusehen, dass auch Wochen nach dem Laborbesuch noch eine geringe, durch den einmaligen Besuch im Labor bedingte, Steigerung des Berufsinteresses bekundet wird. Für zukünftige Studien wäre es sinnvoll, das Berufsinteresse an jedem der drei Messzeitpunkte zu erheben, um tatsächliche Veränderungen und nicht nachträgliche Einschätzungen zu erfassen. In der vorliegenden Studie wurde dies nicht umgesetzt, da das Berufsinteresse lediglich als Nebenaspekt erhoben wurde und der kognitive Wissenstest zu jedem Messzeitpunkt bereits relativ lange Fragebögen bedingte.

Das in Anlehnung an Pawek (2009) aufgestellte Modell bietet vielversprechende Möglichkeiten. So ist es zukünftig denkbar, seine Gültigkeit an unterschiedlichen Gruppen zu testen. Von Interesse wäre insbesondere seine Anwendung auf Datensätze, welche im schulischen Kontext erhoben wurden. So ist anzunehmen, dass in diesem Rahmen die persönliche Disposition eine deutlich größere Rolle spielt, als dies im Schülerlabor der Fall ist. Denkbar sind auch Mehrgruppen-Kausalanalysen (Weiber & Mühlhaus 2010) zwischen den in der vorliegenden Studie untersuchten Subgruppen. Beispielsweise wäre zu prüfen, ob zwischen männlichen und weiblichen Schülern Unterschiede in der Bedeutung der persönlichen Disposition oder in der Wirkung des Laborbesuches bestehen.

4.2.2.6 Selbsteinschätzung von Lehrkräften

Bisher wurden die Ergebnisse von Haupt- und Nebenstudie diskutiert. In diesem Rahmen wurde bereits angemerkt, dass Schüler nur wenig Experimentiererfahrung aus dem Biologieunterricht mitbringen (Scharfenberg 2005, Seidel et al. 2007). Diese Feststellung wurde zum Anlass genommen, um Motive von Lehrkräften für die geringe Experimentierhäufigkeit zu erfragen. Insbesondere wurde der neu entwickelte FSWEx eingesetzt, um folgende Fragestellung zu klären:

Steht die experimentbezogene Selbstwirksamkeitserwartung in Zusammenhang mit der von diesen Lehrkräften präferierten Experimentierform, sowie mit der Experimentierhäufigkeit?

Zur Beantwortung dieser Fragestellung wurden Biologielehrkräfte nach Problemen gefragt, die sie beim Experimentieren in der Schule sehen. Etwa ein Viertel der befragten Lehrpersonen nannte als Problem die eigene Unsicherheit, was sich im Empfinden geringer Selbstwirksamkeit bezogen auf die Planung, Durchführung und Auswertung von Experimenten widerspiegelt. Sicherlich ist nicht auszuschließen, dass einige unsichere Lehrkräfte aufgrund sozialer Erwünschtheit (Bortz & Döring 2006) nicht angeben, dass sie sich unsicher fühlen. Zwar kann dies hier nicht ausgeschlossen werden, als zusätzliche Kontrolle wurde aber die Skala Experimentiersicherheit eingesetzt. Einige ihrer Items erfragen die Experimentiersicherheit indirekt über die Bereitschaft, neue Experimente im Unterricht

auszuprobieren. Die Wahrscheinlichkeit sozial erwünschter Angaben ist in dieser Skala augenscheinlich geringer. Die relativ hohen Korrelation von $r > 0,6$ mit den Subskalen des FSWEx schließen soziale Erwünschtheit zwar nicht aus, deuten aber darauf hin, dass diese eher kein Problem darstellt. Würde dennoch bei einem Teil der unsicheren Lehrkräfte sozial erwünschtes Antwortverhalten vorliegen, so würde sich der beobachtete Unterschied der sicheren und unsicheren Lehrkräfte vermutlich sogar verstärken, wenn diese ehrlich antworten würden. Da nur sehr wenige Lehrkräfte angeben, sehr selten (< 1 x im Schuljahr) oder sehr häufig (1 x in der Woche) zu experimentieren, ist keine sichere Aussage über den Zusammenhang von Experimentiersicherheit und Experimentierhäufigkeit zu belegen. Es deutet sich aber an, dass die Lehrkräfte, die Schülerexperimente präferieren sich kompetenter einschätzen, als die Lehrkräfte, die Lehrerversuche bevorzugen. Diese Beobachtung überrascht zunächst, denn es könnte angenommen werden, dass unsichere Lehrkräfte eher ungern selbst experimentieren. Berücksichtigt man aber, welche Verantwortung ein Lehrer übernimmt, der einen Biologiekurs selbständig experimentieren lässt, ist verständlich, warum die unsicheren Lehrer eher Demonstrationsversuche wählen. Zur Absicherung der Ergebnisse an einer größeren Stichprobe wird diese Untersuchung aktuell weitergeführt.

Gerade hinsichtlich des relativ hohen Anteils von unsicheren Lehrkräften erscheint es wichtig, das Potential von Schülerlaboren auch im Rahmen von Lehrerfortbildungen zu nutzen. Diese können zwar nicht die häufig mäßigen materiellen Bedingungen an den Schulen ändern, aber zu einer nachhaltigen Steigerung der Experimentiersicherheit und -bereitschaft der Lehrkräfte beitragen. Es ist davon auszugehen, dass sich dies auch auf das Experimentierverhalten im schulischen Unterricht und die Häufigkeit von Schülerexperimenten auswirken könnte. So gehen auch Davies et al. (2008) davon aus, dass Lehrer, die im Umgang mit Versuchsmaterialien sicher sind, diese auch häufiger im Unterricht einsetzen werden. Es ist anzunehmen, dass dies nicht nur für den Umgang mit Gerätschaften gilt, sondern in entsprechend angelegten Fortbildungen auch für Planungs- und Auswertungsprozesse von Experimenten. Der Erfolg von Lehrerfortbildungen an Schülerlaboren scheint gesichert, so äußern doch viele Lehrer den Wunsch, Unterstützung in Form von Materialien, Handreichungen und verbesserter Aus- und Fortbildung zu erhalten (Brandt et al. 2003). In vielen Laboren werden diese auch bereits angeboten, doch das Potenzial der Labore in der Lehrerbildung wird bisher noch nicht voll genutzt (Dähnhardt et al. 2009, Euler 2009b). Der FSWEx ist ein nützliches, ökonomisch einsetzbares Messinstrument, welches auch im Rahmen von Lehrerfortbildungen eingesetzt werden kann. Im Pre-Post-Einsatz können durch die Fortbildung bedingte Änderungen der experimentbezogenen Selbstwirksamkeit verfolgt und – falls notwendig – entsprechende Optimierungsmaßnahmen im Fortbildungsprogramm eingeleitet werden.

5 Ausblick

Wie die Ergebnisse der vorliegenden Studie und die Erfahrung aus mehreren Jahren praktischer Schülerlaborarbeit zeigen, liegt der Grund der relativ geringen Experimentierhäufigkeit im schulischen Kontext scheinbar nicht am Unwillen der Schüler. Das Interesse am Experimentieren erweist sich als unabhängig von der Schulnote und sogar als ausgesprochen hoch. Wie die Ergebnisse der vorliegenden Evaluierung zeigen, gibt es sogar eine Gruppe von Schülern, die ein sehr geringes Interesse am Fach Biologie, aber ein hohes Interesse am praktischen Experimentieren haben. Experimentieren – sei es in der Schule oder am außerschulischen Lernort – bietet somit die Chance, auch diese Schüler zu erreichen. In der vorliegenden Studie deutete sich an, dass gerade die fachlich wenig begeisterten, aber am Experimentieren interessierten Praktiker die Lernumgebung Labor ähnlich positiv wahrnehmen wie ihre fachlich interessierten Mitschüler. Deutlichere Unterschiede scheinen dagegen im schulischen Biologieunterricht gegeben zu sein. Dieses Potential von Schülerlaboren, eine heterogene Schülerschaft relativ homogen zu fördern, gilt es in kommenden Studien genauer zu untersuchen. Gezielte Vergleichsstudien zwischen beiden Lernorten sollten dazu beitragen, die durch die Lernumgebung bedingten Einflussfaktoren zu identifizieren. Einen Beitrag zu dieser Differenznivellierung zwischen verschiedenen Schülertypen könnte die deutlich höhere Authentizität der Lernumgebung Schülerlabor leisten. Diese kann jedoch nicht allein durch die Tätigkeit des Experimentierens erreicht werden. Es ist anzunehmen, dass insbesondere Laborkurse mit deutlichem Bezug zur Lebenswelt und unter Berücksichtigung aktueller, alltagsrelevanter Forschung diesen Zweck erfüllen. Umso wichtiger erscheint es, die Wirkung von Schülerlaborkursen nicht losgelöst vom fachwissenschaftlichen Inhalt zu studieren. Viel mehr Wert sollte auf die fachlich fundierte Entwicklung von Kursen sowie deren Optimierung unter didaktischen Gesichtspunkten gelegt werden. Denn naturwissenschaftliche Schülerlabore wollen für Naturwissenschaften begeistern und dies kann nicht ohne fachwissenschaftliche Basis gelingen.

Dass dieser Ansatz im BeLL Bio funktioniert, soll abschließend die Äußerung eines Schülers im Laborkurs Genetischer Fingerabdruck zeigen:

„Endlich hab ich gecheckt, was im Biunterricht dran war.“

6 Literaturverzeichnis

- Abd-Elsalam K. A. (2003). Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology* 2: (5) 91–95.
- acatech deutsche Akademie der Technikwissenschaften (2008). *Monitoring von Motivationskonzepten für den Technikenachwuchs*. Berlin, Heidelberg. Springer Verlag.
- Adl S. M., Simpson A. G. B., Farmer M. A., Andersen R. A., Anderson O. R., Barta J. R., Bowser S. S., Brugerolle G., Fensome R. A., Fredericq S., James T. Y., Karpov S., Kugrens P., Krug J., Lane C. E., Lewis L. A., Lodge J., Lynn D. H., Mann D. G., McCourt R. M., Mendoza L., Moestrup O., Mozley-Standridge S. E., Nerad T. A., Shearer C. A., Smirnov A. V., Spiegel F. W., Taylor M. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* 52: (5) 399-451.
- Adl S. M., Leander B. S., Simpson A. G. B., Archibald J. M., Anderson O. R., Bass D., Bowser S. S., Brugerolle G., Farmer M. A., Karpov S., Kolisko M., Lane C. E., Lodge D. J., Mann D. G., Meisterfeld R., Mendoza L., Moestrup O., Mozley-Standridge S. E., Smirnov A. V., Spiegel F. (2007). Diversity, nomenclature, and taxonomy of protists. *Systematic Biology* 56 (4). 684-689.
- Aepfelbach M. (2011). *Konzeption und Durchführung eines Kurstages im Schülerlabor BeLL Bio zum Thema "Außer Atem - wie kommt der Kater in den Muskel" mit anschließender Evaluation*. BA-Thesis, Bergische Universität Wuppertal.
- Ahmad S., Seshadri M. (2007). Genetic polymorphism of six DNA loci in six population groups of India. *Human Biology* 79: (4) 427–443.
- Albert R., Koster C. J. (2002). *Empirie in Linguistik und Sprachlehrforschung: Ein methodologisches Arbeitsbuch*. Tübingen. Gunter Narr.
- Ammermann C., Gluchowski P., Schmidt P. (1975). Rekursive oder nicht-rekursive Modelle?: Zum Problem der Testbarkeit von Feedback-Prozessen. *Zeitschrift für Soziologie* 4: (3) 203–220.
- Anton M. A., Hock K. (2003). "Uni Meets School": Universität als außerschulischer Lernort. In Pitton A. (Hrsg.). *Ausserschulisches Lernen in Physik und Chemie*. Münster. LIT.
- Arisholm E., Gallis H., Dyba T., Sjöberg D. I. K. (2007). Evaluating Pair Programming with Respect to System Complexity and Programmer Expertise. *IEEE Transactions on Software Engineering* 33: (2) 65-86.
- Asendorpf J. B. (2007). *Psychologie der Persönlichkeit*. Heidelberg: Springer.
- Aufdermauer A., Hesse M. (2006a). Eine Analyse von Biologie-Schulbüchern unter besonderer Berücksichtigung des Experimentierens mit Pflanzen. *IDB Münster* (15) 1–32.
- Aufdermauer A., Hesse M. (2006b). Die Meinung von Schülern zu ihrem Biologie-Schulbuch und zu Schulversuchen. *IDB Münster* (15) 105–112.
- Babushkina N., Kucher A. (2011). Functional role of VNTR polymorphism of human genes. *Russian Journal of Genetics* 47: (6) 637–645.

- Backhaus U., Braun T. (2009). Impulse für offenes Experimentieren: Forschendes Lernen in der Physik. In Messner R. (Hrsg.). *Schule forscht*. Hamburg. Ed. Körber-Stiftung.
- Bagozzi R. P. (1981). Evaluating structural equation models with unobservable variables and measurement error: a comment. *Journal of Marketing Research* 18: (3) 375–381.
- Baldauf S. L. (2003). The deep roots of eukaryotes. *Science* 300: (5626) 1703–1706.
- Baldauf S. L. (2008). An overview of the phylogeny and diversity of eukaryotes. *Journal of Systematics and Evolution* 46: (3) 263–273.
- Ballantyne K. N., Goedbloed M., Fang R. X., Schaap O., Lao O., Wollstein A., Choi Y., van Duijn K., Vermeulen M., Brauer S., Decorte R., Poetsch M., Wurmb-Schwark N. v., Knijff P. d., Labuda D., Vezina H., Knoblauch H., Lessig R., Roewer L., Ploski R., Dobosz T., Henke L., Henke J., Furtado M. R., Kayser M. (2010). Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications. *American Journal of Human Genetics* 87: (3) 341–353.
- Baltes-Götz B. (2010). Analyse von Strukturgleichungsmodellen mit Amos 18. <http://www.uni-trier.de/fileadmin/urt/doku/amos/v18/amos18.pdf>. zuletzt geprüft am 05.01.2012 Trier.
- Bandura A. (1982). Self-Efficacy Mechanism in Human Agency. *American Psychologist* 37: (2) 122–147.
- Barke H.-D., Harsch G. (2011). Alltag und Chemie. In Barke H.-D., Harsch G. (Hrsg.). *Chemiedidaktik kompakt*. Berlin, Heidelberg. Springer-Verlag.
- Barth U., Pfeifer P. (2009). Lehrerfortbildung im Bereich Chemie - eine Chance für die Unterrichtsentwicklung an der Hauptschule. *CHEMKON* 16: (2) 67–73.
- Baumert J. (2000). *Schülerleistungen im internationalen Vergleich: Eine neue Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten*. Berlin. Max-Planck-Inst. für Bildungsforschung.
- Bäumli-Roßnagl M. A. (1981). Das Unterrichtsexperiment im Urteil der Lehrer: Ergebnisse einer Lehrerbefragung zur Durchführung von Experimenten im Grundschulunterricht. *Sachunterricht und Mathematik in der Primarschule* (4) 156–163.
- Behrendt M. (2009). Eine Reise durch Schülerlabore und Science Center. In Messner R. (Hrsg.). *Schule forscht*. Hamburg. Ed. Körber-Stiftung.
- Bell T., Hackl R. (2006). Crash-Tests: Hypothesengeleitet experimentieren: Arbeitsergebnisse aus dem piko-Schulset Schleswig-Holstein. In Pitton A. (Hrsg.). *Lehren und Lernen mit neuen Medien*. Berlin. LIT-Verl.
- Ben-Zvi R., Hofstein A., Samuel D., Kempa R. F. (1977). Modes of instruction in high school chemistry. *Journal of Research in Science Teaching* 14: (5) 433–439.
- Berck K.-H., Graf D. (2010). *Biologiedidaktik: Grundlagen und Methoden*. (4., vollst. überarb Aufl.). Wiebelsheim. Quelle & Meyer.
- Berg J. M., Stryer L., Tymoczko J. L. (2007). *Biochemie*. (6 Aufl.). München. Elsevier Spektrum Akademischer Verlag.

- Biermann H., Schöpe T., Geigenmüller A., Enke M. (2009). Ingenieurmangel in Deutschland – ein Image- und Kommunikationsproblem? In Nagl M., Bargstädt H.-J., Hoffmann M., Müller N. (Hrsg.). *Zukunft Ingenieurwissenschaften - Zukunft Deutschland*. Berlin, Heidelberg. Springer.
- Bockau U. (2006). *Tetrahymena thermophila: Ein neues Expressionssystem zur Produktion humaner Proteine*. Dissertation, Westfälische Wilhelms-Universität Münster.
- Boerwinkle E., Xiong W. J., Fourest E., Chan L. (1989). Rapid typing of tandemly repeated hypervariable loci by the polymerase chain-reaction - Application to the apolipoprotein-B 3' hypervariable region. *Proceedings of the National Academy of Sciences of the United States of America* 86: (1) 212–216.
- Boos-Nünning U. (2005). Zuhause in zwei Sprachen: Sprachkompetenz von Mädchen mit Migrationshintergrund. In Gogolin I., Krüger-Potratz M., Kuhs K., Neumann U., Wittek F. (Hrsg.). *Interkulturelle Bildungsforschung*. Münster. Waxmann.
- Borrmann T. (2008). Laboratory Education in New Zealand. *Eurasia Journal of Mathematics, Science & Technology Education* 4: (4) 327–335.
- Bortz J., Döring N. (2003). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler*. (3. Aufl.). Heidelberg. Springer.
- Bortz J., Döring N. (2006). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler*. (4. Aufl.). Heidelberg. Springer.
- Bortz J., Lienert G. A., Boehnke K. (2008). *Verteilungsfreie Methoden in der Biostatistik: Mit 247 Tabellen*. (3. Aufl.). Heidelberg. Springer.
- Bortz J., Schuster C. (2010). *Statistik für Human- und Sozialwissenschaftler*. (7 Aufl.). Berlin. Springer.
- Bos W., Bensen M., Baumert J., Prenzel M., Selzer C., Walther G. (Hrsg.). (2008). *TIMSS 2007: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster. Waxmann.
- Brandt A., Möller J., Kohse-Höinghaus K. (2003). Das Chemieexperimentierlabor teutolab als außerschulischer Lernort: Ergebnisse einer Evaluationsstudie. In A. Pitton (Hrsg.). *Ausserschulisches Lernen in Physik und Chemie*. Münster. LIT.
- Brandt A. (2005). *Förderung von Motivation und Interesse durch außerschulische Experimentierlabors*. Dissertation. Universität Bielefeld.
- Brandt A., Koch J., Verbeet M. (2007). Land der kleinen Forscher. *Der Spiegel* (49) 84–100.
- Brandt A., Möller J., Kohse-Höinghaus K. (2008). Was bewirken außerschulische Experimentierlabors? *Zeitschrift für Pädagogische Psychologie* 22: (1) 5–12.
- Browne M., Cudeck R. (1993). Alternative ways of assessing equation model fit. In Bollen K.A., Long J.S. (Hrsg.). *Testing structural equation models*. Newbury Park. Sage Publications.
- Budowle B., Chakraborty R., Giusti A. M., Eisenberg A. J., Allen R. C. (1991). Analysis of the VNTR locus DIS80 by the PCR followed by high-resolution page. *American Journal of Human Genetics* 48: (1) 137–144.

- Bühl A. (2010). *SPSS 18: Einführung in die moderne Datenanalyse*. (12. Aufl.). München. Pearson Studium.
- Bundesverband der Schülerlabore e V. Daten zu Schülerlaboren (2012). <http://www.lernort-labor.de/data.php?tl=12>. zuletzt geprüft am 01.02.2012.
- Bundesverband der Schülerlabore e V. Webseite Lernort-Labor (2011). <http://www.lernort-labor.de/>. zuletzt geprüft am 01.04.2011.
- Burtscher M., Faulhaber M., Kornexl E., Nachbauer W. (2005). Cardiorespiratory and metabolic responses during mountain hiking and downhill skiing. *Wiener Medizinische Wochenschrift* 155: (7-8) 129–135.
- Buscemi L., Cucurachi N., Mencarelli R., Sisti B., Tagliabracci A., Ferrara S. D. (1994). PCR typing of the locus D17S30 (YNZ22 VNTR) in an Italian population-sample. *International Journal of Legal Medicine* 106: (4) 200–204.
- Busse I., Preisfeld A. (2002). Unusually expanded SSU ribosomal DNA of primary osmotrophic Euglenids: Molecular Evolution and Phylogenetic Inference. *Journal of Molecular Evolution* 55: (6) 757-767.
- Busse I., Preisfeld A. (2003). First evidence of a group I intron among euglenozoans discovered in the SSU rDNA of *Ploeotia costata* (Euglenozoa). *Protist* 154: 57-69.
- Bybee R. W. (2002). Scientific Literacy - Mythos oder Realität? In Evans R., Gräber W., Koballa T., Nentwig P. (Hrsg.). *Scientific Literacy*. Opladen. Leske + Budrich.
- Byrne B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. (2 Aufl.). New York, London. Routledge.
- Campbell N. A. (2000). *Biologie*. Heidelberg, Berlin, Oxford. Spektrum Akademischer Verlag.
- Cavalier-Smith T. (2002). Chloroplast Evolution: Secondary Symbiogenesis and Multiple Losses. *Current biology* 12: (2) R62-R64.
- Chandler P., Sweller J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction* 8: (4) 293–332.
- Chin W. W. (1998). Issues and opinion on structural equation modeling. *Management information systems quarterly* 22: (1) 7–16.
- Cimer A. (2007). Effective Teaching in Science: A Review of Literature. *Journal of Turkish Science Education* 4: (1) 20–44.
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. (2 Aufl.). Hillsdale N.J. Lawrence Erlbaum Associates.
- Collins K., Gorovsky M. A. (2005). *Tetrahymena thermophila*. *Current Biology* 15: (9) 317–318.
- Conley A. M., Pintrich P. R., Vekiri I., Harrison D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology* 29: (2) 186–204.

- Csikszentmihalyi M., Schiefele U. (1993). Die Qualität des Erlebens und der Prozeß des Lernens. *Zeitschrift für Pädagogik* 39: (2) 207–221.
- Curran P. J., Finch J. F., West S. G. (1996). The Robustness of Test Statistics to Nonnormality and Specification Error in Confirmatory Factor Analysis. *Psychological Methods* 1: (1) 16–29.
- Dahmen S. (2012). *Videostudie zu Experimentierverhalten und Evaluation des Kurses "Dem Täter auf der Spur" im Bergischen Lehr-Lernlabor*. MA-Thesis, Bergische Universität Wuppertal.
- Dähnhardt D., Haupt O., Pawek C. (2009). Neugier wecken, Kompetenzen fördern: Wie Schülerlabore arbeiten. In Dähnhardt D., Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.
- Damerau K. (2010). *Welchen Nutzen hat Vorbereitung?*: Poster zur Tagung Genlabor & Schule V in Braunschweig. Wuppertal. Bergische Universität Wuppertal.
- Das B., Chauhan P. S., Seshadri M. (2002a). Allele frequency distribution of two minisatellite loci (D19S20 and ApoB 3' VNTR) among five distinct population groups of India. *Journal of Forensic Sciences* 47: (1) 229–231.
- Das B., Ghosh A., Chauhan P. S., Seshadri M. (2002b). Genetic polymorphism study at four minisatellite loci (D1S80, D17S5, D19S20, and APOB) among five Indian population groups. *Human Biology* 74: (3) 345–361.
- Das K., Mastana S. S. (2003). Genetic variation at three VNTR loci in three tribal populations of Orissa, India. *Annals of Human Biology* 30: (3) 237–249.
- Davies R. S., Sprangue C. R., New C. M. (2008). Integrating technology into a science classroom. In Sunal D. W., Wright E., Sundberg C. (Hrsg.). *The impact of the laboratory and technology on learning and teaching science K-16*. Charlotte, N.C. IAP/Information Age Pub.
- Deci E. L., Ryan R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry* 11: (4) 227–268.
- Deci E. L., Ryan R. M. (2012). *Intrinsic Motivation Inventory (IMI): Scale Description*. <http://selfdeterminationtheory.org/questionnaires/10-questionnaires/50>, zuletzt geprüft am 10.02.2012.
- Decorte R., Cuppens H., Marynen P., Cassiman J. J. (1990). Rapid detection of hypervariable regions by the polymerase chain-reaction technique. *DNA and Cell Biology* 9: (6) 461–469.
- Deka R., Chakraborty R., Decroo S., Rothhammer F., Barton S. A., Ferrell R. E. (1992a). Characteristics of polymorphism at a VNTR locus 3' to the apolipoprotein B gene in five human populations. *American Journal of Human Genetics* 51: (6) 1325–1333.
- Deka R., DeCroo S., Mei Yu L., Ferrell, R. E. (1992b). Variable number of tandem repeat (VNTR) polymorphism at locus D17S5 (YNZ22) in four ethnically defined human populations. *Human genetics* 90: (1) 86–90.
- Detlefsen M. (2009). GenaU. In Dähnhardt D., Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.

- Deutsches PISA-Konsortium (2000). *Schülerleistungen im internationalen Vergleich: Eine neue Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten*. Berlin. Max-Planck-Institut für Bildungsforschung.
- DGUV (2010). *Stoffliste zur Regel "Unterricht in Schulen mit gefährlichen Stoffen"*. Berlin. DGUV.
- Dinkelaker J., Herrle M. (2009). *Erziehungswissenschaftliche Videographie: Eine Einführung*. (1 Aufl.). Wiesbaden. VS Verlag für Sozialwissenschaften.
- Dreischmeier A. (in prep.). *Classroom vs. Science Lab – Der Einfluss der Lernumgebung auf Motivation und Interesse beim molekulargenetischen Experimentieren* (Arbeitstitel). Dissertation. Universität Bielefeld.
- Duncan G. T., Balamurugan K., Budowle B., Smerick J., Tracey M. L. (1997). Microvariation at the human D1S80 locus. *International Journal of Legal Medicine* 110: (3) 150–154.
- Ebermann R., Elmadfa I. (2011). *Lehrbuch Lebensmittelchemie und Ernährung*. (2. Aufl.). Wien, New York. Springer.
- Eccles J. S., Adler T. F., Futterman R., Goff S. B., Kaczala C. M., Meece J. L., Migdley C. (1983). Expectancies, values and academic behaviors. In Spence J.T. (Hrsg.). *Achievement and achievement motives*. San Francisco. W.H. Freeman.
- Eccles J. S., Wigfield A. (1995). In the mind of achiever: The structure of adolescent's academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin* 21: 215–225.
- Engeln K., Euler M., Prenzel M., Ringelband U. (2003). Lernort Labor - Experimentieren im außerschulischen Rahmen: Erzeugung von situativem Interesse an Schülerexperimentiertagen. In Pitton A. (Hrsg.). *Ausserschulisches Lernen in Physik und Chemie*. Münster. LIT.
- Engeln K. (2004). Schülerlabors. Authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken. Dissertation, Christian-Albrechts-Universität.
- Engeln K., Euler M. (2004). Forschen statt Pauken: Aktives Lernen im Schülerlabor. *Physik Journal* 3: (11) 45–48.
- Engeln K. (2006). Schülerlabore: Interessenförderung bei Mädchen und Jungen. In Pitton A. (Hrsg.). *Lehren und Lernen mit neuen Medien*. Berlin. LIT-Verl.
- Erlich H.A. (1989). Polymerase chain reaction. *Journal of Clinical Immunology* 9: (6) 437–447.
- Euler M. (2001). Lernen durch Experimentieren. In Ringelband U., Prenzel M., Euler M. (Hrsg.). *Lernort Labor*. Kiel. IPN.
- Euler M. (2009a). Schülerlabore in Deutschland: Zum Mehrwert authentischer Lernorte in Forschung und Entwicklung. *Praxis der Naturwissenschaften - Physik in der Schule* 58: (4) 5–9.
- Euler M. (2009b). Schülerlabore: Lernen, forschen und kreative Potenziale entfalten. In Dähnhardt D., Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.

- Ferguson, M. E.; Burow, M. D.; Schulze, S. R.; Bramel, P.J; Paterson, A.H; Kresovich, S.; Mitchell, S. (2004). Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *Theor Appl Genet* 108: (6) 1064–1070.
- Fraser B. J. (1980). Development and validation of a test of enquiry skills. *Journal of Research in Science Teaching* 17: (1) 7–16.
- Freedman M. P. (1997). Relationship among Laboratory Instruction, Attitude toward Science, and Achievement in Science Knowledge. *Journal of Research in Science Teaching* 34: (4) 343–357.
- Frossard P. M., Obineche E. N., Lestringant G. G. (1999). Association of an apolipoprotein B gene marker with essential hypertension. *Hypertension* 33: (4) 1052–1056.
- Fuchs R., Schwarzer R. (1994). Selbstwirksamkeit zur sportlichen Aktivität: Reliabilität und Validität eines neuen Meßinstruments. *Zeitschrift für Differentielle und Diagnostische Psychologie* 15: (3) 141–154.
- Galkin A. S., Grivennikova V. G., Vinogradov A. D. (2001). Stoichiometry of the NADH:Ubiquinone Reductase Reaction Catalyzed by Submitochondrial Particles. *Biochemistry (Moscow)* 66: (4) 435–443.
- Gatti S. (2009). Netzwerk der Schülerlabore in der Helmholtz-Gemeinschaft. In Dähnhardt D. , Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.
- Gebhard U. (2008). Schülerinnen und Schüler. In Gropengießer H. (Hrsg.). *Fachdidaktik Biologie*. Köln. Aulis-Verl. Deubner.
- Gentürk E. (2009). *Experimente im Biologieunterricht: Definition, Bedeutung, Kompetenzerwerb*. München. GRIN Verlag GmbH.
- Glaserfeld E. v. (2010). Konstruktion der Wirklichkeit und des Begriffs der Objektivität. In Gumin H., Meier H. (Hrsg.). *Einführung in den Konstruktivismus*. München. Piper.
- Glowinski I. (2007). *Schülerlabore im Themenbereich Molekularbiologie als Interesse fördernde Lernumgebungen*. Dissertation. Christian-Albrechts-Universität.
- Goldstein P. Z., DeSalle R. (2011). Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description. *Bioessays* 33: (2) 135–147.
- Goodwin W., Linacre A., Hadi S. (2011). *An introduction to forensic genetics*. Chichester, West Sussex, UK. Wiley-Blackwell.
- Graßhoff G., Casties R., Nickelsen K. (2000). *Zur Theorie des Experiments: Untersuchungen am Beispiel der Entdeckung des Harnstoffzyklus*. Bern. Institut für Philosophie.
- Gropengießer H., Kattmann U. (Hrsg.). (2008). *Fachdidaktik Biologie*. Köln. Aulis-Verl. Deubner.
- Groß J., Gropengießer H. (2003). Erfassung von Lernprozessen mittels retrospektiver Befragung in Natur- und Erlebniswelten. *Erkenntnisweg Biologiedidaktik* 2: 91–102.
- Großschedl J., Harms U. (2008). Similarity Judgments Test: Ein Verfahren zur Erfassung von Wissensstrukturen. *Erkenntnisweg Biologiedidaktik* 7: 85–100.

- Grube C. (2011). *Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung*. Dissertation. Universität Kassel.
- Gruehn S. (2000). Unterricht und schulisches Lernen. In Rost D. H. (Hrsg.). *Pädagogische Psychologie und Entwicklungspsychologie*. Münster. Waxmann Verlag.
- Gruenberger C. (2007). *Atomic Force Microscopy of the alga Euglena gracilis*. Diplomarbeit. Technische Universität Wien.
- Guderian P., Priemer B., Schön L. H. (2006a). Die Entwicklung von Interesse bei mehrfachen und in den Unterricht eingebundenen Schülerlaborbesuchen. In Pitton A. (Hrsg.). *Lehren und Lernen mit neuen Medien*. Berlin. LIT-Verl.
- Guderian P., Priemer B., Schön L. H. (2006b). In den Unterricht eingebundene Schülerlaborbesuche und deren Einfluss auf das aktuelle Interesse an Physik. *Physik und Didaktik in Schule und Hochschule* (5/2). 142–149.
- Guderian P. (2007). *Wirksamkeitsanalyse außerschulischer Lernorte*. Dissertation. Humboldt-Universität zu Berlin.
- Gundlach E. (2003). Nach dem PISA-Schock: Höhere Bildungsausgaben oder umfassende Bildungsreform? In Hansel T. (Hrsg.). *PISA -Und die Folgen?*. Herbolzheim. Centaurus-Verl.
- Günther J., Grygier P., Kircher E., Thoermer C., Sodian B. (2003). Epistemologische Überzeugungen von SchülerInnen und LehrerInnen. In Pitton A. (Hrsg.). *Außerschulisches Lernen in Physik und Chemie*. Münster. LIT.
- Hadler M. (2005). *Quantitative Datenanalyse für Sozialwissenschaftler*. Wien. LIT-Verl.
- Hammann M. (2004). *Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung - dargestellt anhand von Kompetenzen beim Experimentieren*. MNU 57: (4) 196–203.
- Hammann M. (2007). Das Scientific Discovery as Dual Search-Modell. In Krüger D., Vogt H. (Hrsg.). *Theorien in der biologiedidaktischen Forschung*. Berlin, New York. Springer.
- Hammann M., Hoi Phan T., Bayrhuber H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? *Zeitschrift für Erziehungswissenschaft* 10: (8) 33–49.
- Hampel P., Kropf V., Dikici S., König L., Gloger-Tippelt G., Petermann, F. (2007). Kognitive Entwicklung und sozial-emotionale Kompetenzen bei Frühgeborenen mit unterschiedlichen Bindungsrepräsentationen. *Kindheit und Entwicklung* 16: (4) 220-228.
- Hascher T., Moser P. (2001). Betreute Praktika - Anforderungen an Praktikumslehrerinnen und -lehrer. *Beiträge zur Lehrerbildung* 19: (2) 217–231.
- Häußler P., Bünder W., Duit R., Gräber W., Mayer J. (1998). *Naturwissenschaftsdidaktische Forschung - Perspektiven für die Unterrichtspraxis*. Kiel. IPN.
- Heidecke G., Kropf J., Stork G. (1983). Quantitative Derivatphotometrie an digitalisierten Spektren. *Fresenius' Zeitschrift für Analytische Chemie* 316: (4) 405–408.

- Helbig M. (2012). Warum bekommen Jungen schlechtere Schulnoten als Mädchen? Ein sozialpsychologischer Erklärungsansatz. *Zeitschrift für Bildungsforschung* 2: (1) 41-54.
- Helms J. V. (1998). Learning about the Dimensions of Science through Authentic Tasks. In Wellington J. (Hrsg.). *Practical work in school science*. London, New York. Routledge.
- Hempelmann R. (2009). SaarLab. In Dähnhardt D. , Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.
- Hetze P. (2011). *Nachhaltige Hochschulstrategien für mehr MINT-Absolventen*. (2. Aufl.). Essen. Stifterverband für die Wissenschaftspflege.
- Heublein U., Schmelzer R., Sommer D. (2008). *Die Entwicklung der Studienabbruchquote an den deutschen Hochschulen: Ergebnisse einer Berechnung des Studienabbruchs auf der Basis des Absolventenjahrgangs 2006*. Hannover. Hochschul-Informations-System GmbH (HIS). <http://www.his.de/pdf/21/his-projektbericht-studienabbruch.pdf>, zuletzt geprüft am 01.03.12.
- Heublein U., Hutzsch C., Schreiber J., Sommer D., Besuch G. (2009). *Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen: Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres 2007/08*. Hannover. Hochschul-Informations-System GmbH (HIS). http://www.his.de/pdf/21/studienabbruch_ursachen.pdf, zuletzt geprüft am 01.03.12.
- Hillebrandt D., Dähnhardt D. (2005). *Forschend lernen – Schülerlabore in Deutschland: Naturwissenschaft und Nachwuchs gehen ein Stück des Weges gemeinsam*. TheoPrax (1) 20–23.
- Hoffmann L., Häußler P., Lehrke M. (1998). *Die IPN-Interessenstudie Physik*. Kiel. IPN.
- Hoffmeister N. (2003). *Klassenfahrten*. München. GRIN Verlag.
- Hofstein A., Lunetta V. N. (2003). The laboratory in science education: Foundations for the twenty-first century. *Science Education* 88: (1) 28–54.
- Holstermann N., Bögeholz S. (2007). Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften* 13: 71–86.
- Homburg C., Baumgartner H. (1995). Beurteilung von Kausalmodellen. Bestandsaufnahme und Anwendungsempfehlungen. *Marketing: Zeitschrift für Forschung und Praxis* 17: (3) 162-176.
- Homburg C., Giering A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte: Ein Leitfaden für die Marketingforschung. *Marketing: Zeitschrift für Forschung und Praxis* 18: (1) 5–24.
- Hopf M., Wiesner H. (2003). Problemlösen beim Experimentieren. In Pitton A. (Hrsg.). *Ausserschulisches Lernen in Physik und Chemie*. Münster. LIT.
- Horn F. (2009). *Biochemie des Menschen: Das Lehrbuch für das Medizinstudium*. Stuttgart, New York. Thieme.

- Horn G.T; Richards B.; Klinger K.W. (1989). Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. *Nucleic Acids Research* 17: (5) 2140.
- Höttecke D. (2001). Die Vorstellungen von Schülern und Schülerinnen von der "Natur der Naturwissenschaften". *Zeitschrift für Didaktik der Naturwissenschaften* 7: 7–23.
- Howe C., Smith P. (1998). Experimentation and Conceptual Understanding in School Science: Can hypothesis testing play a role? In Wellington J. (Hrsg.). *Practical work in school science*. London , New York. Routledge.
- Huck J., Haan G. d., Plesse M. (2010). *Schülerlabor & Co: Außerschulische naturwissenschaftlich-technische Experimentierangebote als Ergänzung des Schulunterrichts in der Region Berlin-Brandenburg*. Berlin. Regioverlag.
- Hwang I. T., Kim Y. J., Kim S. H., Kwak C. I., Gu Y. Y., Chun J. Y. (2003). Annealing control primer system for improving specificity of PCR amplification. *BioTechniques* 35: (6) 1180–1184.
- IPN Kiel (2012). Das Programm Videograph. Kiel. IPN. <http://www.ipn.uni-kiel.de/aktuell/videograph/htmStart.htm>. zuletzt geprüft am 17.03.2012.
- Ivey J. N., Atchison B. A., Georgalis A. M. (1994). Assessment of PCR of the D17S30 Locus for Forensic Identification. *Journal of Forensic Sciences* 39: (1) 52-63.
- Jasinska A., Krzyzosiak W. J. (2004). Repetitive sequences that shape the human transcriptome. *FEBS Letters* (567) 136–141.
- John H., Thinesse-Demel J. (2004). *Lernort Museum – neu verortet!: Ressourcen für soziale Integration und individuelle Entwicklung. Ein europäisches Praxishandbuch*. Bielefeld. Transcript.
- Jung S. (2002). *Forensische DNA-Analytik: Methodische, spurenkundliche und populationsgenetische Betrachtungen unter besonderer Berücksichtigung mitochondrialer DNA-Polymorphismen*. Dissertation. Julius-Maximilians-Universität.
- Kaluza G., Schulze H.-H. (2000). Evaluation von Gesundheitsförderungsprogrammen - Methodische Stolpersteine und pragmatische Empfehlungen. *Zeitschrift für Gesundheitspsychologie* 8: (1) 18–24.
- Kaminska A. (2010). *Entwicklung und Evaluierung eines Kurses für das Bergische Lehr-Lernlabor (BeLL Bio) zum Thema „Kleine Einzeller ganz groß“ unter besonderer Berücksichtigung der Biologie des Ciliaten Tetrahymena thermophila*. BA-Thesis. Bergische Universität Wuppertal.
- Karrer K. M. (1999). Tetrahymena Genetics: Two Nuclei Are Better Than One. In Asai D. J., Forney J. D. (Hrsg.). *Tetrahymena thermophila*. San Diego. Academic Press.
- Kasai K., Nakamura Y., White R. (1990). Amplification of a Variable Number of Tandem Repeats (VNTR) Locus (pMCT118) by the Polymerase Chain Reaction (PCR) and its Application to Forensic Science. *Journal of Forensic Science* 35: (5) 1196–1200.
- Kayser M., Sajantila A. (2001). Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Science International* 118: (2) 116–121.

- Keeling P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365: (1541) 729-748.
- Keuneke S., Graß H., Ritz-Timme S. (2010). „CSI-Effekt“ in der deutschen Rechtsmedizin. *Rechtsmedizin* 20: (5) 400–406.
- Kiechle M. (2009). SchullaborBayern (SLB). In Dähnhardt D., Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.
- Killermann W., Hiering P., Starosta B. (Hrsg.). (2009). *Biologieunterricht heute: Eine moderne Fachdidaktik*. Donauwörth. Auer.
- Kirchner S., Priemer B. (2007). Probleme von Schülern mit offenen Experimentieraufgaben in Physik. In Höttecke D. (Hrsg.). *Naturwissenschaftlicher Unterricht im internationalen Vergleich*. Berlin. LIT.
- Klahr D., Dunbar K. (1988). Dual space search during scientific reasoning. *Cognitive Science* 12: 1–48.
- Klahr D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, Mass. MIT Press.
- Klautke S. (1974). Kriterien zur Beurteilung von Schulbüchern für Biologie. *Blickpunkt Schulbuch* (16) 30–32.
- Klein H. P., Krautz J. (2012). Soll Qualität wirklich durch Notendumping gesichert werden? *FAZ* 15.03.2012: 8.
- Klieme E., Avenarius H., Blum W., Döbrich P., Gruber H., Prenzel M., Reiss K., Riquarts K., Rost J., Tenorth H.-E., Vollmer H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards: Expertise*. Berlin. Bundesministerium für Bildung und Forschung (BMBF).
- Kline R. B. (2011). *Principles and Practice of Structural Equation Modeling*. New York, London. The Guilford Press.
- Kloosterman A. D., Budowle B., Daselaar P. (1993). PCR amplification and detection of the human DIS80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. *International Journal of Legal Medicine* 105: 257-264.
- Klug W. S., Spencer C. A., Cummings M. R. (2007). *Genetik*. München. Pearson Studium.
- KMK (2005). *Bildungsstandards im Fach Biologie für den mittleren Schulabschluss (Jahrgangsstufe 10): Beschluss vom 16.12.2004*. München, Neuwied. Luchterhand.
- Koch H. J., Raschka C. (1999). Emprische und Mechanistische Auswertungsmodelle fuer Laktatkonzentrationszeitkurven im Dauerleistungstest. *Nagoya Journal of Medical Science* 62: 127–134.
- Köhler K. (2004). Welche fachgemäßen Arbeitsweisen werden im Biologieunterricht eingesetzt? In Spörhase-Eichmann U. (Hrsg.). *Biologie-Didaktik*. Berlin. Cornelsen Scriptor.
- Köller O., Trautwein U., Lüdtke O., Baumert J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie* 20: (1) 27–39.

- Köller O., Möller J. (2010). Selbstwirksamkeit. In Rost D. H. (Hrsg.). *Handwörterbuch pädagogische Psychologie*. Weinheim. Beltz.
- Krapp A. (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie in Erziehung und Unterricht* 45: (3) 185–201.
- Krapp A., Ryan R. M. (2002). Selbstwirksamkeit und Lernmotivation. In Jerusalem M., Hopf, D. (Hrsg.). *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen*. Weinheim. Beltz.
- Krapp A. (2010). Interesse. In Rost D. H. (Hrsg.). *Handwörterbuch pädagogische Psychologie*. Weinheim. Beltz.
- Kühl S., Strodtholz P., Taffertshofer A. (Hrsg.). (2009). *Experiment*. Wiesbaden. VS Verlag für Sozialwissenschaften.
- Leach J. (1998). Teaching about the World of Science in the Laboratory. In Wellington J. (Hrsg.). *Practical work in school science*. London , New York. Routledge.
- Leedale G. F. (1967). Euglenoid Flagellates. *Englewood Cliffs, NJ. Prentice-Hall*.
- Leedale G. F. (1978). Phylogenetic criteria in euglenoid flagellates. *Biosystems* 10: (1-2) 183-187.
- Lehrer R., Schauble L., Petrosino A. J. (2001). Reconsidering the role of experiment in science education. In Crowley K., Schunn C., Okada T. (Hrsg.). *Designing for science: Implications from everyday, classroom, and professional settings*. Mahwah, N.J. Lawrence Erlbaum Associates.
- Leibold K. (1997). *Modelle, Modellbildung und Modelleinsatz*. Dissertation. Universität Bayreuth.
- Leisen J. (2006). *Aufgabenkultur im mathematisch-naturwissenschaftlichen Unterricht*. MNU 59: (5) 260–266.
- Lewis B. P., Linder D. E. (1997). Thinking about Choking? Attentional Processes and Paradoxical Performance. *Personality and Social Psychology Bulletin* 23: (9) 937–944.
- Lienert G. A., Raatz U. (1998). *Testaufbau und Testanalyse*. Weinheim. Beltz, Psychologie-Verl.-Union.
- Lim G. S., Balke M., Meier R. (2012). Determining Species Boundaries in a World Full of Rarity: Singletons, Species Delimitation Methods. *Systematic Biology* 61: (1) 165–169.
- Lottspeich F., Engels J. W. (2006). *Bioanalytik*. München, Heidelberg. Elsevier, Spektrum, Akademischer Verlag.
- Löwe B. (1987). Interessenverfall im Biologieunterricht. *Unterricht Biologie* 11: (124) 62–65.
- Ludwig P. H. (2003). Partielle Geschlechtertrennung – enttäuschte Hoffnungen?: Monoedukative Lernumgebungen zum Chancenausgleich im Unterricht auf dem Prüfstand. *Zeitschrift für Pädagogik* 49: (5) 640–656.
- Ma C., Bloomfield V.A. (1995). Gel Electrophoresis Measurement of Counterion Condensation on DNA. *Biopolymers* 35: (2) 211-216.

- Maiseyenko V., Nawrath D., Schecker H. (2011). Modellbasierte Diagnose und Förderung experimenteller Kompetenz. In Höttecke D. (Hrsg.). *Naturwissenschaftliche Bildung als Beitrag zur Gestaltung partizipativer Demokratie Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Potsdam 2010*. Berlin, Münster: LIT.
- Margulis L., Chapman M. J. (2009). *Kingdoms and Domains: Illustrated Phyla of Life: An Illustrated Guide to the Phyla of Life on Earth*. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo. Academic Press.
- Marsh H. W. (2005). Big-Fish-Little-Pond Effect on Academic Self-Concept. *Zeitschrift für Pädagogische Psychologie* 19: (3) 119–129.
- Masters R., Nott M. (1998). Implicit Knowledge and Science Practical Work in Schools. In Wellington J. (Hrsg.). *Practical work in school science*. London, New York. Routledge.
- Mayer J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In Krüger D., Vogt H. (Hrsg.). *Theorien in der biologiedidaktischen Forschung*. Berlin, New York. Springer.
- Mayer J., Grube C., Möller A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In Harms U., Sandmann, A. (Hrsg.). *Lehr- und Lernforschung in der Biologiedidaktik*. Innsbruck, Wien, Bozen. Studienverlag.
- Messner R. (2009). Forschendes Lernen: Impulse zur Klärung fachlicher Schwerpunkte. In Messner R. (Hrsg.). *Schule forscht*. Hamburg. Ed. Körber-Stiftung.
- Meyer H. (2004). *Was ist guter Unterricht?: Studienbuch*. Berlin. Cornelsen Verlag Scriptor.
- Ministerium für Innovation Wissenschaft u. Forschung NRW (2011). Webseite der Initiative Zukunft durch Innovation (zdi). <http://www.wissenschaft.nrw.de/zdi/>. zuletzt geprüft am 29.03.2011.
- Ministerium für Innovation Wissenschaft u. Forschung NRW (2012). zdi-Schülerlabore. <http://www.wissenschaft.nrw.de/zdi/zdi-Schuelerlabore/index.php>. zuletzt geprüft am 01.02.2012.
- Mitchell M. (1993). Situational Interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology* 85: (3) 424–436.
- Mittag W., Jerusalem M. (1999). Determinanten des Rauchverhaltens bei Jugendlichen und Transfereffekte eines schulischen Gesundheitsprogrammes. *Zeitschrift für Gesundheitspsychologie* 7: (4) 183–202.
- Möller A., Grube C., Meyer J., (2007). Kompetenzniveaus der der Erkenntnisgewinnung bei Schülerinnen und Schülern der Sekundarstufe I. In Bayrhuber H., Harms U., Krüger D., Sandmann A., Unterbruner U., Upmeyer zu Belzen A., Vogt H. (Hrsg.). *Ausbildung und Professionalisierung von Lehrkräften*. Kassel. IPN.
- Möller J., Trautwein U. (2009). Selbstkonzept. In E. Wild E., Möller J. (Hrsg.). *Pädagogische Psychologie*. Heidelberg. Springer.
- Morling N. (2009). PCR in forensic genetics. *Biochemical Society Transactions* 37: 438–440.
- Moschner B., Dickhäuser O. (2010). Selbstkonzept. In Rost D. H. (Hrsg.). *Handwörterbuch pädagogische Psychologie*. Weinheim. Beltz.

- MSW NRW (2008a). *Richtlinien und Lehrpläne für die Grundschule in Nordrhein-Westfalen: Deutsch, Sachunterricht, Mathematik, Englisch, Musik, Kunst, Sport, Evangelische Religionslehre, Katholische Religionslehre*. Frechen. Ritterbach.
- MSW NRW (2008b). *Kernlehrplan für das Gymnasium - Sekundarstufe I in Nordrhein-Westfalen - Biologie*. Frechen. Ritterbach.
- MSW NRW. (2011a). *Kernlehrplan für die Hauptschule in Nordrhein-Westfalen - Lernbereich Naturwissenschaften: Biologie, Chemie, Physik*. Frechen. Ritterbach.
- MSW NRW (2011b). *Das Schulwesen in Nordrhein-Westfalen aus quantitativer Sicht 2010/2011*. Düsseldorf. MSW.
- MSW NRW (2011c). *Kernlehrplan für die Realschule in Nordrhein-Westfalen - Biologie*. Frechen. Ritterbach.
- MSWWF NRW (1999). *Richtlinien und Lehrpläne für die Sekundarstufe II, Gymnasium, Gesamtschule in Nordrhein-Westfalen, Biologie*. Frechen. Ritterbach.
- Mülhardt C. (2009). *Der Experimentator: Molekularbiologie, Genomics*. Heidelberg. Spektrum Akad.-Verlag.
- Mullis K., Faloona F., Scharf S., Saiki R., Horn G., Erlich H. (1986). Specific enzymatic amplification of DNA invitro - the polymerase chain-reaction. *Cold Spring Harbor Symposia on Quantitative Biology* 51: 263-273.
- Naito E., Dewa K., Fukuda M., Sumi H., Wakabayashi Y., Umetsu K., Yuasa I., Yamanouchi H. (2003). Novel paternity testing by distinguishing parental alleles at a VNTR locus in the differentially methylated region upstream of the human H19 gene. *Journal of Forensic Sciences* 48: (6) 1275–1279.
- Nakamura Y., Carlson M., Krapcho K., White R. (1988). Isolation and mapping of a polymorphic DNA sequence (pMCT118) on chromosome 1p (D1S80). *Nucleic Acids Research* 16: (19) 9364-9364.
- Nakamura Y., Koyama, K., Matsushima, M. (1998). VNTR (variable number of tandem repeat) sequences as transcriptional, translational, or functional regulators. *Journal of Human Genetics* 43: (3) 149–152.
- Natzke H., Petermann F. (2009). Schulbasierte Prävention aggressiv-oppositionellen und dissozialen Verhaltens: Wirksamkeitsüberprüfung des Verhaltenstrainings für Schulanfänger. *Praxis der Kinderpsychologie und Kinderpsychiatrie* 58: (1) 34–50.
- Nawrath D., Maiseyenka V., Schecker H. (2011). Experimentelle Kompetenz - Ein Modell für die Unterrichtspraxis. *Praxis der Naturwissenschaften - Physik in der Schule* 60: (6) 42–48.
- Nellen W. (2009). Genlabor & Schule. In Dähnhardt D. , Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.
- Neumann, K. (2004). Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker. In Niedderer, H., Fischler, H. (Hrsg.). *Studien zum Physiklernen Band 38*. Berlin. Logos.

- Neumeyer S., Mikelskis H. F. (2006). Lernort Praktikum im Grund- und Leistungskurs der Klassenstufe 11. In Pitton A. (Hrsg.). *Lehren und Lernen mit neuen Medien*. Berlin. LIT-Verl.
- Nick S., Urhahne D. (2004). Helfen 3D-Simulationen beim Lernen mit CHEMnet? *CHEMKON* 11: (1) 27–31.
- Osborne J. (1998). Science Education without a Laboratory. In Wellington J. (Hrsg.). *Practical work in school science*. London , New York. Routledge.
- Ozkan S., Cakiroglu J., Tekkaya C. (2008). Students' Perceptions of the Science Laboratory Learning Environment. In Sunal D. W., Wright E., Sundberg C. (Hrsg.). *The impact of the laboratory and technology on learning and teaching science K-16*. Charlotte, N.C. IAP/Information Age Pub.
- Parchmann I. (2009). Alltagsorientierung in den Naturwissenschaften: Forschendes Lernen im Chemieunterricht. In Messner R. (Hrsg.). *Schule forscht*. Hamburg. Ed. Körber-Stiftung.
- Parmentier M. (1995). Antwort auf die Frage: Was ist Museumspädagogik? *Standbein Spielbein*: (43) 6–7.
- Pawek C. (2009). *Schülerlabore als interessefördernde außerschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe*. Dissertation. Christian-Albrechts-Universität zu Kiel.
- Pfangert-Becker U. (2010). Das Experiment im Lehrund Lernprozess: Eine Betrachtung aus der Sicht kompetenzorientierten Lehrens und Lernens im Kontext der zweiten Ausbildungsphase. *Praxis der Naturwissenschaften - Chemie in der Schule* 59: (6) 40–42.
- Pfister D. (2008). Wie Change- und Raumgestaltungs-Management verbunden werden können. *OrganisationsEntwicklung* (3) 63–73.
- Pfister D. (2011). *Raum - Atmosphäre - Nachhaltigkeit: Emotionale und kulturelle Aspekte der sozialen Nachhaltigkeit des Bauens, des Immobilienmarketings und der Gebäudebewirtschaftung*. Basel. Edition gesowip.
- Pohlmann B., Möller J., Streblow L. (2005). Bedingungen leistungsbezogenen Verhaltens im Sportunterricht. *Zeitschrift für Sportpsychologie* 12: (4) 127–134.
- Pohlmann B., Möller J. (2006). Vergleichseffekte auf kognitive, affektive und motivationale Variablen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 38: (2) 79–87.
- Popper K., Miller D. (1983). A proof of the impossibility of inductive probability. *Nature* 302: (5910) 687–688.
- Preisfeld A., Busse I., Klingberg M., Talke S., Ruppel H. G. (2001). Phylogenetic position and inter-relationships of the osmotrophic euglenids based on SSU rDNA data, with emphasis on the Rhabdomonadales (Euglenozoa). *International Journal of Systematic and Evolutionary Microbiology* 51: (3) 751–758.
- Prenzel M., Ringelband U. (2001). "Lernort Labor" - Neue Initiativen. In Ringelband U., Prenzel M., Euler M. (Hrsg.). *Lernort Labor*. Kiel. IPN.

- Prenzel M. (2007). PISA 2006: Wichtige Ergebnisse im Überblick. In Prenzel M., Artelt C., Baumert J., Blum W., Hammann M., Klieme E., Pekrun R. (Hrsg.). *PISA 2006 - Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster, München: Waxmann.
- Priemer B. (2006). Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen. *Zeitschrift für Didaktik der Naturwissenschaften* 12: 159–175.
- Quatember A. (2011). *Statistik ohne Angst vor Formeln: Das Studienbuch für Wirtschafts- und Sozialwissenschaftler*. München. Pearson Studium.
- Rakoczy K., Klieme E., Pauli C. (2008). Die Bedeutung der wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse und des Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation. *Zeitschrift für Pädagogische Psychologie* 22: (1) 25–35.
- Randler C., Bogner F. X. (2008). Planning Experiments in Science Education Research: Comparison of a Quasi-Experimental Approach with a Matched Pair Tandem Design. *International Journal of Environmental & Science Education* 3: 95–103.
- Rasch B., Frieze M., Hofmann W., Naumann E. (2010a). *Quantitative Methoden 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Berlin, Heidelberg. Springer.
- Rasch B., Frieze M., Hofmann W., Naumann E. (2010b). *Quantitative Methoden 1: Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Berlin, Heidelberg. Springer.
- Reinmann G., Mandl H. (2006). Unterrichten und Lernumgebungen gestalten. In Krapp A. (Hrsg.). *Pädagogische Psychologie*. Weinheim. BeltzPVU.
- Renneberg R. (2009). *Bioanalytik für Einsteiger: Diabetes, Drogen und DNA*. Heidelberg. Spektrum, Akad. Verlag.
- Richter G. (2003). *Praktische Biochemie: Grundlagen und Techniken*. Stuttgart. Thieme.
- Rommel K., Bernt E., Schmitz F., Grimm K. (1968). Enzymatische Galaktosebestimmung im Blut und oraler Galaktose-Toleranztest. *Klinische Wochenschrift* 46: (17) 936–940.
- Rönnebeck S., Schöps K., Prenzel M., Mildner D., Hochweber J. (2010). Naturwissenschaftliche Kompetenz von PISA 2006 bis PISA 2009. In Klieme E., Artelt C., Hartig J., Jude N., Köller O., Prenzel M., Schneider W., Stanat P. (Hrsg.). *PISA 2009*. Münster, Westf. Waxmann.
- Rost J. (2004). *Lehrbuch Testtheorie - Testkonstruktion*. Bern. Huber.
- Roth G. (2009). Die Bedeutung von Motivation und Emotion für den Lernerfolg. In Messner R. (Hrsg.). *Schule forscht*. Hamburg. Ed. Körber-Stiftung.
- Roth W.-M. (Hrsg.). (1995). *Authentic school science: Knowing and Learning in Open-Inquiry Science Laboratories*. Dordrecht, Boston, London: Kluwer Academic Publishers.
- Ruangjirachuporn W., Japa O., Tanratanavijit M., Nanakorn S. (2006). Population genetic data on D1S80, D17S5, ApoB, COL2A1 and Ig-JH in Northeastern Thais. *Legal Medicine* 8: (5) 308–309.
- Rückmann I. (2009). ForschNet. In Dähnhardt D., Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.

- Ryan R. M., Deci E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* 25: 54–67.
- Sadava D., Hillis D.M., Heller H.C., Berenbaum M.R (2011). *Purves Biologie*. Heidelberg. Spektrum Akad. Verlag.
- Sajantila A., Budowle B., Ström M., Johnsson V., Lukka M., Peltonen L., Ehnholm C. (1992). PCR amplification of alleles at the D1S80 Locus: Comparison of a finnish and a north american caucasian population sample, and forensic casework evaluation. *Am. J. Hum. Genet.*: (50) 816–825.
- Sambrock J., Fritsch E. F., Maniatis T. (1989). *Molecular cloning. A laboratory manual*. New York. Cold Spring Harbor Laboratory Press.
- Sandoval W. A. (2008). Design Principles for Effective Laboratory Instruction. In Sunal D. W., Wright E., Sundberg C. (Hrsg.). *The impact of the laboratory and technology on learning and teaching science K-16*. Charlotte, N.C.. IAP/Information Age Pub.
- Satow L. (2000). *Klassenklima und Selbstwirksamkeitsentwicklung*. Dissertation. Freie Universität Berlin.
- Saur P., Joneleit M., Tölke H., Pudel V., Niedmann P. D., Kettler D. (2002). Evaluation des Magnesiumstatus bei Ausdauersportlern. *Deutsche Zeitschrift für Sportmedizin* 53: (3) 72–78.
- Scharfenberg F. J. (2005). *Experimenteller Biologieunterricht zu Aspekten der Gentechnik im Lernort Labor: empirische Untersuchung zu Akzeptanz, Wissenserwerb und Interesse*. Dissertation. Universität Bayreuth.
- Schecker H., Parchmann I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften* 12: 45–66.
- Schiefele U., Krapp A., Schreyer I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 25: (2) 120–148.
- Schiefele U., Streblow L. (2006). Motivation aktivieren. In Mandl H., Friedrich H. F. (Hrsg.). *Handbuch Lernstrategien*. Göttingen. Hogrefe.
- Schmidt M., Kunkel M., Schuff-Werner P., Naumann M., Reichmann H., Reimers C. D. (1997). Standardisierte aerobe Gehbelastung auf dem Laufband bei Gesunden sowie Patienten mit mitochondrialen und nichtmitochondrialen Myopathien. *Der Nervenarzt* 68: (10) 831–835.
- Schön M. (2007). Medizinische Psychologie und Soziologie. In Priewe J., Tümmers D. (Hrsg.). *Das Erste - Kompendium Vorklinik*. Heidelberg. Springer.
- Schreiber N., Theyßen H., Schecker H. (2009). Experimentelle Kompetenz messen?! *Physik und Didaktik in Schule und Hochschule* 8: (3) 92–101.
- Schulte K. (2008). *Selbstwirksamkeitserwartungen in der Lehrerbildung*. Dissertation. Georg-August-Universität.
- Schulz von Thun F. (2011). *Miteinander reden*. Reinbek bei Hamburg. Rowohlt Taschenbuch-Verlag.

- Seidel T., Prenzel M., Wittwer J., Schwindt K. (2007). Unterricht in den Naturwissenschaften. In PISA-Konsortium Deutschland (Hrsg.). *PISA 2006*. Münster, München. Waxmann.
- Shavelson R. J., Ruiz-Primo M. A., Wiley E. W. (1999). Note on Sources of Sampling Variability in Science Performance Assessments. *Journal of Educational Measurement* 36: (1) 61–71.
- Silva A. J., Johnson J. P., White R. L. (1987). Characterization of a highly polymorphic region 5' to JH in the human immunoglobulin heavy chain. *Nucleic Acids Research* 15: (9) 3845-3857.
- Silverthorn D. U. (2009). *Physiologie*. München. Pearson Studium.
- Singer-Sam J., Tanguay R. L., Riggs A. D. (1989). Use of chelex to improve the PCR signal from a small number of cells. *Amplifications* 3: 11.
- Smeitink J., Wendel U. (2007). Mitochondriopathien. In Lentze M. J., Schulte F. J., Schaub J., Spranger J. (Hrsg.). *Pädiatrie*. Berlin, Heidelberg. Springer Berlin Heidelberg.
- Stanat P., Artelt C., Baumert J., Klieme E., Neubrand M., Prenzel M., Schiefele U., Schneider W., Schümer G., Tillmann K. J., Weiß M. (2003). PISA und PISA-E: Zusammenfassung der bereits vorliegenden Befunde. In Baumert J., Artelt C., Klieme E., Neubrand M., Prenzel M., Schiefele U., Schneider W., Tillmann K. J., Weiß M. (Hrsg.). *PISA 2000, ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen. Leske + Budrich.
- Standop J. (2002). *Emotionen und kognitives schulisches Lernen aus interdisziplinärer Perspektive: Emotionspsychologische, neurobiologische und schulpädagogische Zusammenhänge - ihre Berücksichtigung im schulischen Bildungsauftrag wie den Forschungen zum Unterrichtsklima und der Klassenführung*. Frankfurt am Main. Lang.
- Sunal D. W., Sunal C. S., Sundberg C., Wright E. L. (2008). The Importance of Laboratory Work and Technology in Science Teaching. In Sunal D. W., Wright E., Sundberg C. (Hrsg.). *The impact of the laboratory and technology on learning and teaching science K-16*. Charlotte, N.C. IAP/Information Age Pub.
- Tamir P. (1974). An inquiry oriented laboratory examination. *Journal of Educational Measurement* 11: (1) 25–33.
- Tannenbaum R. S. (1971). The development of the test of science processes. *Journal of Research in Science Teaching* 8: (2) 123–136.
- Teifel-Greding J. (2000). DNA-Analysen in der Forensik. *Praxis der Naturwissenschaften-Biologie in der Schule* 49: (2) 1–2.
- Tesch, M., Duit, R. (2004). Experimentieren im Physikunterricht – Ergebnisse einer Videostudie. *Zeitschrift für Didaktik der Naturwissenschaften* 10: 51–69.
- Trautwein U., Lüdtke O. (2005). The Big-Fish-Little-Pond Effect. *Zeitschrift für Pädagogische Psychologie* 19: (3) 137–140.
- Urhahne D., Hopf M. (2004). Epistemologische Überzeugungen in den Naturwissenschaften und ihre Zusammenhänge mit Motivation, Selbstkonzept und Lernstrategien. *Zeitschrift für Didaktik der Naturwissenschaften* 10: 71–87.

- Urhahne D. (2006). Die Bedeutung domänenspezifischer epistemologischer Überzeugungen für Motivation, Selbstkonzept und Lernstrategien von Studierenden. *Zeitschrift für Pädagogische Psychologie* 20: (3) 189–198.
- Vogt H. (2007). Theorie des Interesses und des Nicht-Interesses. In Krüger D., Vogt H. (Hrsg.). *Theorien in der biologiedidaktischen Forschung*. Berlin, New York. Springer.
- Wahl P., Bloch W., Mester J. (2009). Moderne Betrachtungsweisen des Laktats: Laktat ein überschätztes und zugleich unterschätztes Molekül. *Schweizerische Zeitschrift für Sportmedizin und Sporttraumatologie* 57: (3) 100–107.
- Walsh P. S., Metzger D. A., Higuchi R. (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* 10: (4) 506–513.
- Watzl B., Briviba K., Rechkemmer G. (2002). Anthocyane. *Ernährungs-Umschau* 49: (4) 148-150.
- Watzlawick P. (2011). *Wie wirklich ist die Wirklichkeit?*. München. Piper.
- Weber A. P. M., Osteryoung K. W. (2010). From Endosymbiosis to Synthetic Photosynthetic Life. *Plant Physiology* 154: (2) 593-597.
- Wehlen C. (2007). Wissenschaftsverständnis von Schülerinnen und Schülern und Möglichkeiten zur Veränderung. In Langner R. (Hrsg.). *Ordnungen des Denkens*. Berlin, Münster: LIT.
- Weiber R., Mülhhaus D. (2010). *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS*. Berlin, Heidelberg. Springer.
- Wein M. (2009). Die Kraft der Rapsfelder. In Dähnhardt D., Haupt O.J., Pawek C. (Hrsg.). *Kursbuch 2010*. Marburg. Tectum-Verlag.
- Weinert F. E. (2001). *Leistungsmessungen in Schulen*. Weinheim. Beltz-Verl.
- Wellington J. (1998). Practical Work in Science: time for a re-appraisal. In Wellington J. (Hrsg.). *Practical work in school science*. London, New York. Routledge.
- Wiegand P., Bajanowski T., Brinkmann B. (1993). DNA typing of debris from fingernails. *International Journal of Legal Medicine* 106: (2) 81–83.
- Wilde M. (2004). *Biologieunterricht im Naturkundemuseum im Spannungsfeld zwischen Instruktion und Konstruktion*. Dissertation. Universität Bayreuth.
- Wilde M., Bätz K., Kovaleva A., Urhahne D. (2009). Überprüfung einer Kurzskaala intrinsischer Motivation (KIM): Testing a short scale of intrinsic motivation. *Zeitschrift für Didaktik der Naturwissenschaften* 15: 31–45.
- Wittwer, C.T; Fillmore, G. C.; Hillyard, D.R (1989). Automated polymerase chain reaction in capillary tubes with hot air. *Nucleic Acids Research* 17: (11) 4353–4357.
- Wodzinski R. (2002). Mädchen im Physikunterricht. In Kircher E., Schneider W. B. (Hrsg.). *Physikdidaktik in der Praxis*. Heidelberg. Springer.

- Wolff R. K., Nakamura Y., White R. (1988). Molecular characterization of a spontaneously generated new allele at a VNTR locus: No exchange of flanking DNA sequence. *Genomics* 3: (4) 347–351.
- Woolfolk A. (2008). *Pädagogische Psychologie*. München, Boston. Pearson Studium.
- Yakovchuk P., Protozanova E., Frank-Kamenetskii M. D. (2006). Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Research* 34: (2) 564-574.
- Zehren W. (2009). *Forschendes Experimentieren im Schülerlabor*. Dissertation. Universität des Saarlandes.
- Zhang C. S., Xing D. (2009). Parallel DNA amplification by convective polymerase chain reaction with various annealing temperatures on a thermal gradient device. *Analytical Biochemistry* 387: (1) 102–112.
- Zöfel P. (2002). *Statistik verstehen: Ein Begleitbuch zur computergestützten Anwendung*. München. Addison-Wesley.
- Zwick M. M., Renn O. (2000). *Die Attraktivität von technischen und ingenieurwissenschaftlichen Fächern bei der Studien- und Berufswahl junger Frauen und Männer: Eine Präsentation der Akademie für Technikfolgenabschätzung in Baden-Württemberg*. Stuttgart. Akademie für Technikfolgenabschätzung in Baden-Württemberg.

7 Abbildungsverzeichnis

- Abbildung 1: Prozentuale Zusammensetzung des menschlichen Genoms; farblich markiert sind Mini- und Mikrosatelliten (Verändert nach Jasinska & Krzyzosiak 2004)
- Abbildung 2: Nukleotidsequenz eines D1S80 Abschnittes (Datenquelle: Genbank AB121710.1, Kasai et al. 1990)
- Abbildung 3: Der Weg von der DNA zum Genetischen Fingerabdruck (Eigene Abbildung)
- Abbildung 4: Aerober und anaerober Weg des Pyruvats nach der Glykolyse; stark vereinfacht (Eigene Abbildung, in Anlehnung an Silverthorn 2009)
- Abbildung 5: Energiegewinn im aeroben Stoffwechsel (Eigene Abbildung in Orientierung an den Angaben zur ATP-Produktion aus Silverthorn 2009)
- Abbildung 6: Reaktion, welche von der Laktatdehydrogenase (LDH) katalysiert wird (Eigene Abbildung)
- Abbildung 7: Schematischer Aufbau eines Photometers (Eigene Abbildung, in Anlehnung an Richter 2003)
- Abbildung 8: Photometrisch aufgenommenes Absorptionsspektrum von NADH und NAD⁺ im Wellenlängenbereich von 240-400 nm (Verändert nach: <http://www.sigmaaldrich.com/technical-documents/articles/analytix/enzymatic-assay-kits.html>, geprüft am 02.02.2012)

Abbildung 9: Schematischer Stammbaum der Eukaryoten auf der Basis von SSUrDNA-Analysen (Verändert nach Baldauf 2008)

Abbildung 10: Raster-Elektronenmikroskopische Aufnahme von *Tetrahymena thermophila*. (<http://www.miami.muohio.edu/news/media/1588.jpg>, geprüft am 23.03.12)

Abbildung 11: Konjugation bei *Tetrahymena* (Bockau 2006)

Abbildung 12: Lichtmikroskopische Aufnahme von *Euglena gracilis* a) im Licht kultiviert. b) im Dunkeln unter Zugabe von Kohlenstoffquellen kultiviert. (Abbildung mit freundlicher Genehmigung von Prof'in Dr. A. Preisfeld, Bergische Universität Wuppertal)

Abbildung 13: Transmissionselektronenmikroskopische Aufnahme der Pellikula von *Euglenida*. (Abbildung mit freundlicher Genehmigung von Prof'in Dr. A. Preisfeld, Bergische Universität Wuppertal)

Abbildung 14: Das Schülerlabor BeLL Bio liegt in seiner konzeptionellen Offenheit auf einem Kontinuum zwischen dem formellen Charakter der Schule und dem eher informellen Charakter des traditionellen Museums (Abbildung in Anlehnung an Guderian 2007).

Abbildung 15: Einfluss internaler und externaler Vergleiche auf das Fähigkeitsselbstkonzept und Auswirkungen auf Erfolgserwartung, empfundenen Aufgabenwert und Verhalten (Abbildung in Anlehnung an das Erwartungs-Wert-Modell von Eccles et al. 1983).

Abbildung 16: Konzepte von Interesse (Verändert nach Engeln 2004)

Abbildung 17: Das Experiment im naturwissenschaftlichen Lehr-/Lernprozess (Abbildung aus Pfangert-Becker 2010).

Abbildung 18: Die drei Hauptkomponenten des SDDS-Modells (Verändert nach Klahr 1988 & Hammann et al. 2007)

Abbildung 19: Kompetenzmodell naturwissenschaftlichen Denkens (nach Mayer 2007, Mayer et al. 2008)

Abbildung 20: Modell experimenteller Kompetenz (nach Schreiber et al 2009)

Abbildung 21: Teildimensionen experimenteller Kompetenz im Spinnenmodell (Nawrath et al. 2011)

Abbildung 22: Hypothesennetz zentraler affektiver Einflussgrößen im BeLL Bio (Eigene Abbildung)

Abbildung 23: Versuchsdesign Hauptstudie (Eigene Abbildung)

Abbildung 24: Versuchsdesign, der im Rahmen der Hauptstudie durchgeführten Videoanalyse (Eigene Abbildung)

Abbildung 25: Versuchsdesign der Nebenstudie (Eigene Abbildung)

Abbildung 26: Frage zur Erfassung des Migrationshintergrundes (Eigene Abbildung)

Abbildung 27: Frage zur Erfassung der beruflichen Orientierung (Eigene Abbildung)

Abbildung 28: Position der beiden Videokameras im Laborraum des BeLL Bio (Eigene Abbildung)

Abbildung 29: Frage zur Erfassung der Probleme des Experimentierens in der Schule im Lehrerfragebogen (Eigene Abbildung)

Abbildung 30: Schüler beim Auftragen von DNA-Proben auf ein Agarosegel (Eigene Abbildung)

- Abbildung 31: Schülerinnen bei der Arbeit am Photometer im BeLL Bio-Kurs LDH (Eigene Abbildung)
- Abbildung 32: Schüler beim Mikroskopieren von *Tetrahymena thermophila* im BeLL Bio-Kurs Einzeller (Eigene Abbildung)
- Abbildung 33: Pipettierschema der Paarungstypen von *T. thermophila* in der Multiwellplatte (Kaminska 2010)
- Abbildung 34: Die Interpretation des Korrelationskoeffizienten (Verändert nach Quatember 2011)
- Abbildung 35: Pfaddiagramm für ein Strukturgleichungsmodell: Messmodelle (Confirmatorische Faktorenanalysen, CFA) und Strukturmodell sind gekennzeichnet (Abbildung in Anlehnung an Byrne 2010)
- Abbildung 36: Die Bedeutung direkter, indirekter und totaler Effekte im Strukturgleichungsmodell (Eigene Abbildung)
- Abbildung 37: Gelelektrophoretische Auftrennung der PCR-Produkte des VNTR-Locus D1S80 nach verschiedenen DNA-Extraktionsmethoden (Eigene Abbildung)
- Abbildung 38: Gelelektrophoretische Auftrennung (Agarose 1,5% w/v) des Markers 100 bp-DNA-Ladder equimolar mit verschiedenen Färbemethoden (Eigene Abbildung)
- Abbildung 39: Gel-Elektrophoretische Auftrennung der PCR-Produkte nach Gradienten-PCR mit dem Primerpaar für den Locus D17S80 (Eigene Abbildung)
- Abbildung 40: Gelelektrophoretische Aufnahme (Agarose 1,5% w/v) der untersuchten VNTR-Loci nach PCR mit Proben von Spendern unterschiedlicher ethnischer Herkunft (Eigene Abbildung)
- Abbildung 41: Gelelektrophoretische Aufnahme der PCR-Produkte aller vier VNTR-Loci für ME (Spuren 1, 3, 5, 7) und Syr (Spuren 2, 4, 6, 8) (Eigene Abbildung)
- Abbildung 42: Absorptionsspektrum der Erfrischungsgetränke (Eigene Abbildung)
- Abbildung 43: Spektrum der Absorptionsänderung in Probe A (Eigene Abbildung)
- Abbildung 44: Spektrum der Absorptionsänderung in Probe B (Eigene Abbildung)
- Abbildung 45: Spektrum der Absorptionsänderung in Probe C (Eigene Abbildung)
- Abbildung 46: Abhängigkeit des Interesses am Unterrichtsfach Biologie (Fachinteresse) und am Experimentieren (Sachinteresse) von der Schulnote (Eigene Abbildung)
- Abbildung 47: Mittelwerte und Standardabweichungen der Schüler im Fach- und Sachinteresse in Abhängigkeit von der beruflichen Orientierung (Eigene Abbildung)
- Abbildung 48: Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 für weibliche und männliche Schüler (Eigene Abbildung)
- Abbildung 49: Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 für Schüler mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung (Eigene Abbildung)
- Abbildung 50: Durchschnittliche Höhe des empfundenen Beitrages von Experimenten zum Verständnis von Schülern mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung einschließlich der Standardabweichungen (Eigene Abbildung)

- Abbildung 51: Mittelwerte und Standardabweichungen des Wissenschaftsverständnisses für Schüler mit naturwissenschaftlicher und nicht naturwissenschaftlicher beruflicher Orientierung (Eigene Abbildung)
- Abbildung 52: Vorwissen im Wissenspretest Genetischer Fingerabdruck differenziert nach Geschlecht inkl. Standardabweichung (Eigene Abbildung)
- Abbildung 53: Vorwissen im Wissenspretest LDH differenziert nach Geschlecht inkl. Standardabweichung (Eigene Abbildung)
- Abbildung 54: Vorwissen im Wissenspretest Einzeller differenziert nach Geschlecht inkl. Standardabweichung (Eigene Abbildung)
- Abbildung 55: Vorwissen im Wissenspretest Genetischer Fingerabdruck differenziert nach beruflicher Orientierung inkl. Standardabweichung (Eigene Abbildung)
- Abbildung 56: Vorwissen im Wissenspretest LDH differenziert nach beruflicher Orientierung inkl. Standardabweichung (Eigene Abbildung)
- Abbildung 57: Mittelwerte in den affektiven dispositionalen Personenvariablen der drei aus der Clusteranalyse hervorgegangenen Gruppen Allrounder, Praktiker und Theoretiker (Eigene Abbildung)
- Abbildung 58: A-D: Mittelwerte in den affektiven dispositionalen Personenvariablen der drei aus der Clusteranalyse hervorgegangenen Gruppen Allrounder, Praktiker und Theoretiker für die Versuchsgruppe Genetischer Fingerabdruck (A), die Versuchsgruppe LDH (B), die Versuchsgruppe Einzeller (C) und die Kontrollgruppe (D) (Eigene Abbildung)
- Abbildung 59: Durchschnittliche Höhe der Einschätzungen der Schüler zur Wahrnehmung der affektiven dispositionalen Variablen im Pretest in den drei Laborkursen sowie in der schulischen Kontrollgruppe inklusive Standardabweichungen (Eigene Abbildung)
- Abbildung 60: Durchschnittliches Vorwissen der Treatmentgruppe Gen. Fingerabdruck sowie der entsprechenden Kontrollgruppe (Eigene Abbildung)
- Abbildung 61: Durchschnittliches Vorwissen der Treatmentgruppe LDH sowie der entsprechenden Kontrollgruppe (Eigene Abbildung)
- Abbildung 62: Durchschnittliches Vorwissen der Treatmentgruppe Gen. Fingerabdruck sowie der entsprechenden Kontrollgruppe (Eigene Abbildung)
- Abbildung 63: Durchschnittliche Zustimmung der Schüler zur Wahrnehmung der Laborvariablen in den drei Laborkursen sowie in der schulischen Kontrollgruppe inklusive Standardabweichungen. Die Variable Anleitung wurde in der Kontrollgruppe nicht erhoben (Eigene Abbildung)
- Abbildung 64: Mittelwerte in den Laborvariablen der drei aus der Clusteranalyse hervorgegangenen Gruppen Allrounder, Praktiker und Theoretiker für die Versuchsgruppe Genetischer Fingerabdruck (A), die Versuchsgruppe LDH (B), die Versuchsgruppe Einzeller (C) und die Kontrollgruppe (D) (Eigene Abbildung)
- Abbildung 65: Ergebnisse des Wissenstests im Laborkurs Genetischer Fingerabdruck für die Versuchs- und Kontrollgruppe (Eigene Abbildung)
- Abbildung 66: Ergebnisse des Wissenstests zum Laborkurs Genetischer Fingerabdruck für die Schülertypen Allrounder, Praktiker und Theoretiker (Eigene Abbildung)

Abbildung 67: Ergebnisse des Wissenstests im Laborkurs LDH für die Versuchs- und Kontrollgruppe (Eigene Abbildung)

Abbildung 68: Ergebnisse des Wissenstests im Laborkurs Einzeller für die Versuchs- und Kontrollgruppe (Eigene Abbildung)

Abbildung 69: Entwicklung des Fähigkeitsselbstkonzeptes im Pre-, Post- und Follow-up-Test für die drei Versuchsgruppen und die Kontrollgruppe (Eigene Abbildung)

Abbildung 70: Entwicklung des Fähigkeitsselbstkonzeptes der drei Schülertypen Allrounder, Praktiker und Theoretiker über die drei Messzeitpunkte in der Versuchsgruppe Genetischer Fingerabdruck (Eigene Abbildung)

Abbildung 71 Entwicklung des Fähigkeitsselbstkonzeptes der drei Schülertypen Allrounder, Praktiker und Theoretiker über die drei Messzeitpunkte in der Versuchsgruppe LDH (Eigene Abbildung)

Abbildung 72: Durchschnittliche Höhe der drei Komponenten des aktuellen Interesses im Post- und Follow-up-Test einschließlich der Standardabweichungen (Eigene Abbildung)

Abbildung 73: Mittelwerte der FSWEx.V1-Subskala Planung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 74: Mittelwerte der postFSWEx.V1-Subskala Planung der Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller im Posttest inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 75: Mittelwerte der FSWEx.V1-Subskala Durchführung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 76: Mittelwerte der FSWEx.V1-Subskala Durchführung der Allrounder, Praktiker und Theoretiker in Pre- und Follow-up-Test der Versuchsgruppe Genetischer Fingerabdruck inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 77: Mittelwerte der FSWEx.V1-Subskala Durchführung der Allrounder, Praktiker und Theoretiker in Pre- und Follow-up-Test der Versuchsgruppe LDH inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 78: Mittelwerte der postFSWEx.V1-Subskala Durchführung der Versuchsgruppen Genetischer Fingerabdruck, LDH und Einzeller im Posttest inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 79: Mittelwerte der FSWEx.V1-Subskala Auswertung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 80: Mittelwerte der FSWEx.V1-Subskala Auswertung der Treatmentgruppen in Pre- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)

Abbildung 81: Mittelwerte und Standardabweichung der Einschätzung, wie sehr der BeLL Bio Besuch das Schülerinteresse für naturwissenschaftliche Berufe geweckt hat (Eigene Abbildung)

Abbildung 82: Mittelwerte und Standardabweichung der Einschätzung, wie sehr der BeLL Bio-Kurs Genetischer Fingerabdruck das Interesse für naturwissenschaftliche Berufe geweckt hat (Eigene Abbildung)

- Abbildung 83: Mittelwerte und Standardabweichung der Einschätzung, wie sehr der BeLL Bio-Kurs LDH das Interesse für naturwissenschaftliche Berufe geweckt hat (Eigene Abbildung)
- Abbildung 84: Strukturgleichungsmodell zum Einfluss eintägiger BeLL Bio-Besuche auf die experimentbezogene Selbstwirksamkeitserwartung (orange), das aktuelle Interesse (violett), das biologische Fähigkeitsselbstkonzept (oliv) und das Berufsinteresse (schwarz, Eigene Abbildung)
- Abbildung 85: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Planung der Versuchs- und Kontrollgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)
- Abbildung 86: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Planung der männlichen und weiblichen Schüler der Versuchsgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)
- Abbildung 87: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Durchführung der Versuchs- und Kontrollgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)
- Abbildung 88: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Auswertung der Versuchs- und Kontrollgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)
- Abbildung 89: Ergebnisse der Nebenstudie - Mittelwerte der FSWEx.V2-Subskala Auswertung der männlichen und weiblichen Schüler der Versuchsgruppe in Pre-, Post- und Follow-up-Test inklusive Standardabweichungen (Eigene Abbildung)
- Abbildung 90: Äußerungen von Lehrern zu Problemen beim Experimentieren am Lernort Schule. (Eigene Abbildung)
- Abbildung 91: Mittelwerte und Standardabweichungen der selbstsicheren und unsicheren Lehrpersonen in der subjektiv empfundenen Selbstsicherheit (Eigene Abbildung)
- Abbildung 92: Grafische Darstellung der Mittelwerte und Standardabweichungen der selbstsicheren und unsicheren Lehrpersonen in den Selbstwirksamkeits-Subskalen des FSWEx.V1 (Eigene Abbildung)
- Abbildung 93: Grafische Darstellung der Mittelwerte und Standardabweichungen der selbstsicheren und unsicheren Lehrpersonen im wertebezogenen Interesse an Schülerlaborkursen (Eigene Abbildung)
- Abbildung 94: Grafische Darstellung der Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 in Abhängigkeit von der Experimentierhäufigkeit der Lehrpersonen im Unterricht (Eigene Abbildung)
- Abbildung 95: Grafische Darstellung der Mittelwerte und Standardabweichungen der Selbstwirksamkeits-Subskalen des FSWEx.V1 für Lehrpersonen, die häufiger Lehrerversuche bzw. Schülerversuche durchführen bzw. durchführen lassen (Eigene Abbildung)

Danksagung

An erster Stelle möchte ich mich bei Frau Prof'in Dr. Gela Preisfeld für die Möglichkeit bedanken, im Rahmen des Bergischen Lehr-Lernlabors der Biologie dieses spannende Thema bearbeiten zu dürfen. Ihre offene Tür und anregende Diskussionen sorgten für angenehmes und motivierendes Arbeiten.

Bei Herrn Prof. Dr. Norbert Grotjohann möchte ich mich für die Erstellung des Zweitgutachtens bedanken und dafür, dass ich über ihn zum ersten Mal vom Bergischen Lehr-Lernlabor der Biologie erfahren habe.

Dr. Sabrina Bleidißel möchte ich für Rat und ständige Unterstützung, sowie für gelegentliche Grillabende danken.

Frau Sabine Stratmann gebührt Dank für ihre ständige Ansprechbarkeit bei laborpraktischen Problemen.

Weiterhin bedanke ich mich bei unserer gesamten Arbeitsgruppe für die ständige Unterstützung im Schülerlabor, ein hervorragendes Arbeitsklima, das nächtliche Korrekturlesen und...ach was rede ich...Ihr wisst schon wofür!

Bedanken möchte ich mich auch bei Frau Agnes Dreischmeier für die Rund-um-die-Uhr-Betreuung in der Endphase dieser Arbeit und den überregionalen Schülerlaboraustausch.

Mein besonderer Dank gilt meinen Eltern, ohne die mir das Studium und somit diese Arbeit kaum möglich gewesen wäre.

Anhang

A	Umformulierung der affektiven Items im Posttest für die Kontrollgruppe.....	315 -
B	Deskriptive Statistik.....	318 -
B.1	Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest – differenziert nach Schülertypen	318 -
B.2	Deskriptive Statistik der Laborvariablen im Posttest – differenziert nach Schülertypus -	320 -
C	Modell affektiver Einflussgrößen im BeLL Bio	322 -
C.1	Modell affektiver Einflussgrößen am Lernort BeLL Bio – Direkte Effekte	322 -
C.2	Modell affektiver Einflussgrößen am Lernort BeLL Bio – Indirekte Effekte	323 -
C.3	Modell affektiver Einflussgrößen am Lernort BeLL-Bio – Totale Effekte	324 -

A Umformulierung der affektiven Items im Posttest für die Kontrollgruppe

Laborvariablen im Posttest

Tabelle 1: Messinstrument: Alltagsbezug in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Ich konnte eine Bedeutung der durchgeführten Experimente für das alltägliche Leben erkennen.
Kontrollgruppe	Ich kann eine Bedeutung der im Biologieunterricht durchgeführten Experimente für das alltägliche Leben erkennen.
Treatmentgruppe	Ich habe heute etwas über die Bedeutung der Biologie für unseren Alltag gelernt.
Kontrollgruppe	Ich lerne im Biologieunterricht etwas über die Bedeutung der Biologie für unseren Alltag.
Treatmentgruppe	Ich habe heute etwas über die Bedeutung der Biologie für unsere Gesellschaft gelernt.
Kontrollgruppe	Im Biologieunterricht lerne ich etwas über die Bedeutung der Biologie für unsere Gesellschaft.

Tabelle 2: Messinstrument Authentizität in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Ich habe heute etwas darüber gelernt, wie Wissenschaftler arbeiten.
Kontrollgruppe	Im Biologieunterricht lerne ich etwas darüber, wie Wissenschaftler arbeiten.
Treatmentgruppe	Ich habe heute ein Gefühl dafür bekommen, wie Forschung funktioniert.
Kontrollgruppe	Im Biologieunterricht bekomme ich ein Gefühl dafür, wie Forschung funktioniert.
Treatmentgruppe	Ich habe heute einen Einblick in den Berufsalltag von Wissenschaftlern bekommen.
Kontrollgruppe	Ich bekomme im Biologieunterricht einen Einblick in den Berufsalltag von Wissenschaftlern.

Tabelle 3: Messinstrument: Betreuer in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Ich habe das Gefühl, dass die Betreuer von Naturwissenschaften/Technik fasziniert sind.
Kontrollgruppe	Ich habe das Gefühl, dass mein Biologielehrer von Naturwissenschaften/Technik fasziniert ist.
Treatmentgruppe	Ich hatte die Möglichkeit, den Betreuern des BeLL Bio Fragen zu stellen.
Kontrollgruppe	Beim Experimentieren im Biologieunterricht habe ich die Möglichkeit, dem Biologielehrer Fragen zu stellen.
Treatmentgruppe	Die Gespräche mit den Betreuern waren mir wichtig.
Kontrollgruppe	Beim Experimentieren im Biologieunterricht sind mir die Gespräche mit dem Biologielehrer wichtig.

Tabelle 4: Messinstrument: Offenheit in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Während des Experimentierens hatte ich das Gefühl, nichts selber bestimmen zu können.
Kontrollgruppe	Während des Experimentierens im Biologieunterricht habe ich das Gefühl, nichts selber bestimmen zu können.
Treatmentgruppe	Während des Experimentierens hatte ich die Möglichkeit, eigene Ideen auszuprobieren.
Kontrollgruppe	Während des Experimentierens im Biologieunterricht habe ich die Möglichkeit, eigene Ideen auszuprobieren.
Treatmentgruppe	Während der Experimente konnte ich eigene Entscheidungen treffen.
Kontrollgruppe	Im Biologieunterricht kann ich während des Experimentierens eigene Entscheidungen treffen.

Tabelle 5: Messinstrument: Zusammenarbeit in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Ich habe während des Experimentierens mit meinen Mitschülerinnen und Mitschülern über das Experiment diskutiert.
Kontrollgruppe	Während des Experimentierens im Biologieunterricht diskutiere ich mit meinen Mitschülerinnen und Mitschülern über die Experimente.
Treatmentgruppe	Ich habe während der Experimente meinen Mitschülerinnen/Mitschülern etwas erklärt oder mir ist etwas erklärt worden.
Kontrollgruppe	Während des Experimentierens im Biologieunterricht erkläre ich meinen Mitschülerinnen/Mitschülern etwas oder lasse mir etwas erklären.
Treatmentgruppe	Ich habe während des Experimentierens gut mit meinen Mitschülerinnen und Mitschülern im Team zusammengearbeitet.
Kontrollgruppe	Ich arbeite während des Experimentierens im Biologieunterricht gut mit meinen Mitschülerinnen und Mitschülern im Team zusammen.

Tabelle 6: Messinstrument: Druck in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Während der Laborarbeit fühlte ich mich unter Leistungsdruck gesetzt.
Kontrollgruppe	Während des Experimentierens im Biologieunterricht fühle ich mich unter Leistungsdruck gesetzt.
Treatmentgruppe	Ich fühle mich bei der Laborarbeit angespannt.
Kontrollgruppe	Ich fühle mich beim Experimentieren im Biologieunterricht angespannt.

Tabelle 10: Messinstrument: Epistemische Komponente des aktuellen Interesses in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Ich werde in Büchern/im Internet nachlesen, um mehr Informationen über das im BeLL Bio behandelte Thema zu erhalten.
Kontrollgruppe	Ich lese in Büchern/im Internet nach, um mehr Informationen über die im Biologieunterricht behandelten Themen zu erhalten.
Treatmentgruppe	Ich werde außerhalb des Unterrichts über Dinge nachdenken, die wir im BeLL Bio gesehen oder angesprochen haben.
Kontrollgruppe	Ich denke außerhalb des Unterrichts über Dinge nach, die wir im Biologieunterricht gesehen oder angesprochen haben.
Treatmentgruppe	Ich werde mit Freunden, Eltern oder Geschwistern über Dinge sprechen, die ich im BeLL Bio erlebt habe.
Kontrollgruppe	Ich spreche mit Freunden, Eltern oder Geschwistern über Dinge, die ich im Biologieunterricht erlebe.

Aktuelles Interesse im Follow-up-Test

Tabelle 11: Messinstrument: Emotionale Komponente des aktuellen Interesses in Treatment- und Kontrollgruppe (Messzeitpunkt: Follow-up-Test)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Die Experimente am Tag des BeLL Bio Besuches haben Spaß gemacht.
Kontrollgruppe	Die Experimente im Biologieunterricht machen Spaß.
Treatmentgruppe	Die Durchführung der Experimente im BeLL Bio war langweilig.
Kontrollgruppe	Die Durchführung der Experimente im Biologieunterricht ist langweilig.
Treatmentgruppe	Ich fand die Experimente im BeLL Bio spannend.
Kontrollgruppe	Ich finde die Experimente im Biologieunterricht spannend.

Tabelle 12: Messinstrument: Wertebezogene Komponente des aktuellen Interesses in Treatment- und Kontrollgruppe (Messzeitpunkt: Follow-up-Test)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Die im BeLL Bio durchgeführten Experimente erscheinen mir wichtig.
Kontrollgruppe	Die im Biologieunterricht durchgeführten Experimente erscheinen mir wichtig.
Treatmentgruppe	Dass wir im BeLL Bio Experimente durchgeführt haben, erscheint mir sinnvoll.
Kontrollgruppe	Dass wir im Biologieunterricht Experimente durchführen, erscheint mir sinnvoll.

Tabelle 7: Messinstrument: Laboratmosphäre (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Der Laborraum wirkte beunruhigend auf mich.
Kontrollgruppe	Der Biologieraum wirkt beunruhigend auf mich.
Treatmentgruppe	Der Laborraum verbreitete eine angenehme Atmosphäre.
Kontrollgruppe	Der Biologieraum verbreitet eine angenehme Atmosphäre.

Aktuelles Interesse im Posttest

Tabelle 8: Messinstrument: Emotionale Komponente des aktuellen Interesses in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Die Experimente haben Spaß gemacht.
Kontrollgruppe	Die Experimente im Biologieunterricht machen Spaß.
Treatmentgruppe	Die Durchführung der Experimente war langweilig.
Kontrollgruppe	Die Durchführung der Experimente im Biologieunterricht ist langweilig.
Treatmentgruppe	Ich fand die Experimente spannend.
Kontrollgruppe	Ich finde die Experimente im Biologieunterricht spannend.

Tabelle 9: Messinstrument: Wertebezogene Komponente des aktuellen Interesses in Treatment- und Kontrollgruppe (Messzeitpunkt: Posttest)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Die heute durchgeführten Experimente erscheinen mir wichtig.
Kontrollgruppe	Die im Biologieunterricht durchgeführten Experimente erscheinen mir wichtig.
Treatmentgruppe	Dass wir heute Experimente durchgeführt haben, erscheint mir sinnvoll.
Kontrollgruppe	Dass wir im Biologieunterricht Experimente durchführen, erscheint mir sinnvoll.
Treatmentgruppe	Dass wir heute Experimente durchgeführt haben, ist mir persönlich wichtig.
Kontrollgruppe	Dass wir im Biologieunterricht Experimente durchführen, ist mir persönlich wichtig.

Treatmentgruppe	Dass wir im BeLL Bio Experimente durchgeführt haben, ist mir persönlich wichtig.
Kontrollgruppe	Dass wir im Biologieunterricht Experimente durchführen, ist mir persönlich wichtig.

Tabelle 13: Messinstrument: Epistemische Komponente des aktuellen Interesses in Treatment- und Kontrollgruppe (Messzeitpunkt: Follow-up-Test)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Ich habe nach dem BeLL Bio-Besuch in Büchern/im Internet nachgelesen, um mehr Informationen über das im BeLL Bio behandelte Thema zu erhalten.
Kontrollgruppe	Ich lese in Büchern/im Internet nach, um mehr Informationen über die im Biologieunterricht behandelten Themen zu erhalten.
Treatmentgruppe	Ich habe nach dem BeLL Bio-Besuch außerhalb des Unterrichts über Dinge nachgedacht, die wir im BeLL Bio gesehen oder angesprochen haben.
Kontrollgruppe	Ich denke außerhalb des Unterrichts über Dinge nach, die wir im Biologieunterricht gesehen oder angesprochen haben.
Treatmentgruppe	Ich habe nach dem BeLL Bio-Besuch mit Freunden, Eltern oder Geschwistern über Dinge gesprochen, die ich im BeLL Bio erlebt habe.
Kontrollgruppe	Ich spreche mit Freunden, Eltern oder Geschwistern über Dinge, die ich im Biologieunterricht erlebe.

Berufsinteresse

Tabelle 14: Messinstrument: Berufsinteresse (Messzeitpunkt: Follow-up-Test)

Messinstrument: Authentizität	Item-Formulierung
Treatmentgruppe	Nach dem BeLL Bio-Besuch kann ich mir eher vorstellen, eine naturwissenschaftliche Ausbildung/ein naturwissenschaftliches Studium anzutreten, als vor dem Besuch.
Kontrollgruppe	Durch den Biologieunterricht kann ich mir eher vorstellen, eine naturwissenschaftliche Ausbildung/ein naturwissenschaftliches Studium anzutreten.
Treatmentgruppe	Das BeLL Bio hat mein Interesse für naturwissenschaftliche Berufe geweckt.
Kontrollgruppe	Der Biologieunterricht hat mein Interesse für naturwissenschaftliche Berufe geweckt.
Treatmentgruppe	Der BeLL Bio-Besuch hat mich angeregt, mich über naturwissenschaftliche Berufe zu informieren.
Kontrollgruppe	Der Biologieunterricht hat mich angeregt, mich über naturwissenschaftliche Berufe zu informieren.

B Deskriptive Statistik

B.1 Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest – differenziert nach Schülertypen

Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest-
Versuchsgruppe LDH

	Complete Linkage	Mittelwert	Standardabweichung	N
preFSK	Allrounder	3,3523	,62791	88
	Praktiker	1,5926	,98090	63
	Theoretiker	2,9910	,69160	37
	Gesamt	2,6915	1,10660	188
preFachint	Allrounder	3,3011	,55986	88
	Praktiker	1,5873	,72701	63
	Theoretiker	2,2568	1,10316	37
	Gesamt	2,5213	1,07204	188
preSachint	Allrounder	3,2330	,67359	88
	Praktiker	3,0635	,75931	63
	Theoretiker	2,0000	,88976	37
	Gesamt	2,9335	,88032	188
preVerständ	Allrounder	2,8712	,64610	88
	Praktiker	2,2593	,84972	63
	Theoretiker	1,8288	,72700	37
	Gesamt	2,4610	,84103	188
preWissVerständ	Allrounder	2,9744	,59214	88
	Praktiker	2,5317	,66371	63
	Theoretiker	2,4527	,59465	37
	Gesamt	2,7234	,65862	188
preFSWEXV1Plan	Allrounder	2,2083	,62578	88
	Praktiker	1,3862	,63699	63
	Theoretiker	1,3964	,53770	37
	Gesamt	1,7730	,73486	188
preFSWEXV1Durch	Allrounder	3,0568	,68829	88
	Praktiker	2,5714	,85599	63
	Theoretiker	2,4189	,75013	37
	Gesamt	2,7686	,80501	188
preFSWEXV1Ausw	Allrounder	2,6439	,66436	88
	Praktiker	1,9312	,74092	63
	Theoretiker	2,3604	,55209	37
	Gesamt	2,3493	,73878	188

Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest-
Versuchsgruppe Genetischer Fingerabdruck

	Complete Linkage	Mittelwert	Standardabweichung	N
preFSK	Allrounder	3,2692	,64542	104
	Praktiker	1,8188	1,00297	46
	Theoretiker	2,8824	,63550	68
	Gesamt	2,8425	,91702	218
preFachint	Allrounder	3,2788	,61069	104
	Praktiker	1,8261	,72432	46
	Theoretiker	2,5294	,65153	68
	Gesamt	2,7385	,86428	218
preSachint	Allrounder	3,2644	,67175	104
	Praktiker	2,9891	,65395	46
	Theoretiker	2,1397	,72197	68
	Gesamt	2,8555	,84160	218
preVerständ	Allrounder	2,9071	,58285	104
	Praktiker	2,5217	,72573	46
	Theoretiker	2,1078	,62145	68
	Gesamt	2,5765	,71522	218
preWissVerständ	Allrounder	3,0000	,58914	104
	Praktiker	2,6685	,60794	46
	Theoretiker	2,6324	,51193	68
	Gesamt	2,8154	,59474	218
preFSWEXV1Plan	Allrounder	2,0096	,59747	104
	Praktiker	1,3551	,55955	46
	Theoretiker	1,3186	,47293	68
	Gesamt	1,6560	,64670	218
preFSWEXV1Durch	Allrounder	2,9327	,74083	104
	Praktiker	2,5326	,83933	46
	Theoretiker	2,1618	,62558	68
	Gesamt	2,6078	,80097	218
preFSWEXV1Ausw	Allrounder	2,5449	,73304	104
	Praktiker	1,8696	,57726	46
	Theoretiker	2,2353	,47859	68
	Gesamt	2,3058	,68130	218

**Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest-
Versuchsgruppe Einzeller**

	Complete Linkage	Mittelwert	Standardabw eichung	N
prefSK	Allrounder	3,0427	,86240	39
	Praktiker	1,9259	1,04510	18
	Theoretiker	2,9231	,66880	13
	Gesamt	2,7333	,99370	70
prefFachint	Allrounder	3,3974	,47544	39
	Praktiker	1,7222	,80845	18
	Theoretiker	2,2308	,92681	13
	Gesamt	2,7500	,99909	70
preSachint	Allrounder	3,4615	,51790	39
	Praktiker	3,0278	,86555	18
	Theoretiker	2,4231	,60712	13
	Gesamt	3,1571	,74475	70
preVerständ	Allrounder	3,1453	,50026	39
	Praktiker	2,3889	,68837	18
	Theoretiker	2,3590	,46071	13
	Gesamt	2,8048	,66302	70
preWissVerständ	Allrounder	3,1538	,53991	39
	Praktiker	2,2917	,56393	18
	Theoretiker	2,5577	,42271	13
	Gesamt	2,8214	,64710	70
prefSWExV1Plan	Allrounder	2,1538	,62054	39
	Praktiker	1,3148	,57704	18
	Theoretiker	1,3846	,60624	13
	Gesamt	1,7952	,72297	70
prefSWExV1Durch	Allrounder	2,9487	,66684	39
	Praktiker	2,8056	,80693	18
	Theoretiker	2,5769	,44936	13
	Gesamt	2,8429	,67857	70
prefSWExV1Ausw	Allrounder	2,3419	,54337	39
	Praktiker	1,8333	,67881	18
	Theoretiker	2,2051	,56990	13
	Gesamt	2,1857	,61552	70

**Deskriptive Statistik der affektiven dispositionalen Variablen im Pretest -
Kontrollgruppe**

	Complete Linkage	Mittelwert	Standardabw eichung	N
prefSK	Allrounder	3,3422	,68658	75
	Praktiker	1,9872	,75118	26
	Theoretiker	2,9877	,66975	27
	Gesamt	2,9922	,86990	128
prefFachint	Allrounder	3,5400	,55605	75
	Praktiker	2,1731	,63154	26
	Theoretiker	2,9074	,58895	27
	Gesamt	3,1289	,79180	128
preSachint	Allrounder	3,3400	,55847	75
	Praktiker	3,1923	,64926	26
	Theoretiker	2,1296	,74152	27
	Gesamt	3,0547	,78177	128
preVerständ	Allrounder	2,8800	,69671	75
	Praktiker	2,6410	,68588	26
	Theoretiker	2,2346	,54549	27
	Gesamt	2,6953	,70869	128
preWissVerständ	Allrounder	3,0033	,64265	75
	Praktiker	2,6346	,50612	26
	Theoretiker	2,5926	,51956	27
	Gesamt	2,8418	,61946	128
prefSWExV1Plan	Allrounder	2,0133	,69195	75
	Praktiker	1,5513	,58075	26
	Theoretiker	1,4074	,41688	27
	Gesamt	1,7917	,67320	128
prefSWExV1Durch	Allrounder	2,9267	,68622	75
	Praktiker	2,6346	,70082	26
	Theoretiker	2,2222	,56045	27
	Gesamt	2,7188	,71747	128
prefSWExV1Ausw	Allrounder	2,5644	,66435	75
	Praktiker	2,0897	,58544	26
	Theoretiker	2,0494	,30246	27
	Gesamt	2,3594	,63523	128

B.2 Deskriptive Statistik der Laborvariablen im Posttest – differenziert nach Schülertypus

Deskriptive Statistik der Laborvariablen im Posttest – Versuchsgruppe LDH

	Complete Linkage	Mittelwert	Standardabweichung	N
postAlltag	Allrounder	2,5736	,93536	86
	Praktiker	1,6339	,92837	61
	Theoretiker	1,9714	,97139	35
	Gesamt	2,1429	1,02712	182
postAuthent	Allrounder	3,1705	,61939	86
	Praktiker	2,8525	,61320	61
	Theoretiker	2,9810	,67127	35
	Gesamt	3,0275	,64025	182
postBetreuer	Allrounder	2,8953	,76741	86
	Praktiker	2,6230	,95156	61
	Theoretiker	2,5714	,77784	35
	Gesamt	2,7418	,84377	182
postZusammen	Allrounder	2,8915	,82358	86
	Praktiker	2,4863	,63705	61
	Theoretiker	2,6095	,61296	35
	Gesamt	2,7015	,74701	182
postDruck	Allrounder	,4767	,63202	86
	Praktiker	,8607	,87622	61
	Theoretiker	,5714	,60807	35
	Gesamt	,6236	,73571	182
postOffen	Allrounder	1,3140	,91087	86
	Praktiker	1,4508	,78901	61
	Theoretiker	1,2143	,68905	35
	Gesamt	1,3407	,83229	182
postAtmosphäre	Allrounder	3,0698	,82306	86
	Praktiker	2,4262	,84575	61
	Theoretiker	2,5714	,81478	35
	Gesamt	2,7582	,87747	182

Deskriptive Statistik der Laborvariablen im Posttest – Versuchsgruppe Genetischer Fingerabdruck

	Complete Linkage	Mittelwert	Standardabweichung	N
postAlltag	Allrounder	2,9412	,77134	102
	Praktiker	2,4961	,85873	43
	Theoretiker	2,5920	,80783	67
	Gesamt	2,7406	,82118	212
postAuthent	Allrounder	3,1307	,60098	102
	Praktiker	2,9922	,63616	43
	Theoretiker	2,8657	,76993	67
	Gesamt	3,0189	,67269	212
postBetreuer	Allrounder	2,8235	,84913	102
	Praktiker	2,5116	,96046	43
	Theoretiker	2,5224	,84132	67
	Gesamt	2,6651	,87952	212
postZusammen	Allrounder	2,7908	,74450	102
	Praktiker	2,6512	,69371	43
	Theoretiker	2,6169	,74818	67
	Gesamt	2,7075	,73676	212
postDruck	Allrounder	,5735	,73591	102
	Praktiker	,7907	,65676	43
	Theoretiker	,5746	,53125	67
	Gesamt	,6179	,66408	212
postOffen	Allrounder	1,1912	,96227	102
	Praktiker	,9070	,77349	43
	Theoretiker	1,0000	,80716	67
	Gesamt	1,0731	,88327	212
postAtmosphäre	Allrounder	2,8529	,93763	102
	Praktiker	2,6977	,88734	43
	Theoretiker	2,6866	,67888	67
	Gesamt	2,7689	,85341	212

Deskriptive Statistik der Laborvariablen im Posttest – Kontrollgruppe

	Complete Linkage	Mittelwert	Standardabweichung	N
postAlltag	Allrounder	2,5399	,72892	71
	Praktiker	2,3067	,49926	25
	Theoretiker	2,1304	,67971	23
	Gesamt	2,4118	,69258	119
postAuthent	Allrounder	1,9906	,89792	71
	Praktiker	1,7600	,83622	25
	Theoretiker	1,7971	,57506	23
	Gesamt	1,9048	,83265	119
postBetreuer	Allrounder	2,6338	,89824	71
	Praktiker	1,8000	,86603	25
	Theoretiker	2,3043	1,01957	23
	Gesamt	2,3950	,96742	119
postZusammen	Allrounder	2,9437	,67843	71
	Praktiker	2,4800	,71414	25
	Theoretiker	2,4058	,59422	23
	Gesamt	2,7423	,70968	119
postDruck	Allrounder	,7465	,65396	71
	Praktiker	1,2800	,77835	25
	Theoretiker	1,4130	,86145	23
	Gesamt	,9874	,77695	119
postOffen	Allrounder	1,7817	,83552	71
	Praktiker	1,5600	,89350	25
	Theoretiker	1,6957	,97397	23
	Gesamt	1,7185	,87246	119
postAtmosphäre	Allrounder	2,2254	,97390	71
	Praktiker	1,6800	1,06927	25
	Theoretiker	2,2174	,90235	23
	Gesamt	2,1092	,99822	119

Deskriptive Statistik der Laborvariablen im Posttest – Versuchsgruppe
Einzeller

	Complete Linkage	Mittelwert	Standardabweichung	N
postAlltag	Allrounder	2,2523	,90064	37
	Praktiker	1,8431	,79160	17
	Theoretiker	1,8718	1,12660	13
	Gesamt	2,0746	,92990	67
postAuthent	Allrounder	3,0180	,64284	37
	Praktiker	2,8431	,67821	17
	Theoretiker	2,8462	,84564	13
	Gesamt	2,9403	,68884	67
postBetreuer	Allrounder	3,0541	,74334	37
	Praktiker	2,6471	1,27187	17
	Theoretiker	2,6923	1,18213	13
	Gesamt	2,8806	,99274	67
postZusammen	Allrounder	2,9550	,62934	37
	Praktiker	2,8235	,68837	17
	Theoretiker	2,6154	,76795	13
	Gesamt	2,8557	,67476	67
postDruck	Allrounder	,8108	,76695	37
	Praktiker	,7059	,73013	17
	Theoretiker	,8077	,92508	13
	Gesamt	,7836	,77959	67
postOffen	Allrounder	1,7297	1,12790	37
	Praktiker	1,5588	,86390	17
	Theoretiker	1,7308	1,12944	13
	Gesamt	1,6866	1,05470	67
postAtmosphäre	Allrounder	2,8649	,91779	37
	Praktiker	2,6471	,99632	17
	Theoretiker	2,6154	,86972	13
	Gesamt	2,7612	,92260	67

C Modell affektiver Einflussgrößen im BeLL Bio

C.1 Modell affektiver Einflussgrößen am Lernort BeLL Bio – Direkte Effekte

Standardisierte Direkte Effekte (Strukturgleichungsmodell BeLL Bio; bei den Variablen handelt es sich – wenn nicht explizit Einzelitems genannt sind – um Summenscores)

	preSelbstExp	preFSK	Sachinteresse	Fachinteresse	BELLBio	postAktInteresse	postFSK	postSelbstExp	folAktInteresse	folFSK	folSelbstExp	Berufsinteresse
Sachinteresse	,299	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
Fachinteresse	,000	,658	,299	,000	,000	,000	,000	,000	,000	,000	,000	,000
BELLBio	,172	-,120	,102	,304	,000	,000	,000	,000	,000	,000	,000	,000
postAktInteresse	,000	,000	,162	,011	,899	,000	,000	,000	,000	,000	,000	,000
postFSK	,000	,802	,000	,196	,000	,000	,000	,000	,000	,000	,000	,000
postSelbstExp	,546	,000	,000	,000	,437	,000	,000	,000	,000	,000	,000	,000
folAktInteresse	,000	,000	,140	,085	-,164	,724	,000	,000	,000	,000	,000	,000
folFSK	,000	,564	,000	,000	,141	,000	,277	,000	,000	,000	,000	,000
folSelbstExp	,561	,000	,000	,000	,051	,000	,000	,330	,000	,000	,000	,000
Berufsinteresse	,000	,000	,000	,000	,006	,000	,000	,000	,537	,211	,033	,000
folEmo	,000	,000	,000	,000	,000	,000	,000	,000	,722	,000	,000	,000
folWert	,000	,000	,000	,000	,000	,000	,000	,000	,846	,000	,000	,000
folEpi	,000	,000	,000	,000	,000	,000	,000	,000	,721	,000	,000	,000
postEpi	,000	,000	,000	,000	,000	,720	,000	,000	,000	,000	,000	,000
postWert	,000	,000	,000	,000	,000	,847	,000	,000	,000	,000	,000	,000
postEmo	,000	,000	,000	,000	,000	,727	,000	,000	,000	,000	,000	,000
Transparenz	,000	,000	,000	,000	,253	,000	,000	,000	,000	,000	,000	,000
pra20	,000	,000	,000	,939	,000	,000	,000	,000	,000	,000	,000	,000
pra4	,000	,000	,000	,891	,000	,000	,000	,000	,000	,000	,000	,000
Atmosphäre (Einzelitem)	,000	,000	,000	,000	,486	,000	,000	,000	,000	,000	,000	,000
Betreuer (Einzelitem)	,000	,000	,000	,000	,506	,000	,000	,000	,000	,000	,000	,000
pra6	,000	,000	,829	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra12	,000	,000	,442	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra23	,000	,000	,648	,000	,000	,000	,000	,000	,000	,000	,000	,000
foa19	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,796
foa15	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,916
foa1	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,747
foa11	,000	,000	,000	,000	,000	,000	,000	,000	,000	,873	,000	,000
foa28	,000	,000	,000	,000	,000	,000	,000	,000	,000	,937	,000	,000
foa23	,000	,000	,000	,000	,000	,000	,000	,000	,000	,696	,000	,000
poa45	,000	,000	,000	,000	,000	,000	,941	,000	,000	,000	,000	,000
poa41	,000	,000	,000	,000	,000	,000	,904	,000	,000	,000	,000	,000
poa26	,000	,000	,000	,000	,000	,000	,690	,000	,000	,000	,000	,000
pra21	,000	,931	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra14	,000	,893	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra1	,000	,579	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
folPlan	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
folDurch	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,634	,000
folAusw	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,639	,000
posPlan	,000	,000	,000	,000	,000	,000	,000	,479	,000	,000	,823	,000
posDurch	,000	,000	,000	,000	,000	,000	,000	,598	,000	,000	,000	,000
posAusw	,000	,000	,000	,000	,000	,000	,000	,709	,000	,000	,000	,000
prePlan	,644	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
preDurch	,520	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
preAusw	,802	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
Authentizität	,000	,000	,000	,000	,671	,000	,000	,000	,000	,000	,000	,000
Zusammenarb.	,000	,000	,000	,000	,599	,000	,000	,000	,000	,000	,000	,000
Alltagsbezug	,000	,000	,000	,000	,681	,000	,000	,000	,000	,000	,000	,000

C.2 Modell affektiver Einflussgrößen am Lernort BeLL Bio – Indirekte Effekte

Standardisierte Indirekte Effekte (Strukturgleichungsmodell BeLL Bio; bei den Variablen handelt es sich – wenn nicht explizit Einzelitems genannt sind – um Summenscores)

	preSelbstExp	preFSK	Sachinteresse	Fachinteresse	BELLBio	postAktInteresse	postFSK	postSelbstExp	foiAktInteresse	foiFSK	foiSelbstExp	Berufsinteresse
Sachinteresse	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
Fachinteresse	,090	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
BELLBio	,058	,200	,091	,000	,000	,000	,000	,000	,000	,000	,000	,000
postAktInteresse	,256	,079	,177	,273	,000	,177	,000	,273	,000	,000	,000	,000
postFSK	,045	,016	,038	,059	,000	,000	,000	,000	,000	,000	,000	,000
postSelbstExp	,101	,035	,084	,133	,000	,000	,000	,000	,000	,000	,000	,000
foiAktInteresse	,197	,100	,240	,156	,651	,000	,000	,000	,000	,000	,000	,000
foiFSK	,045	,238	,038	,059	,054	,000	,000	,000	,000	,000	,000	,000
foiSelbstExp	,225	,016	,038	,059	,144	,000	,000	,000	,000	,000	,000	,000
Berufsinteresse	,143	,224	,214	,146	,309	,389	,058	,011	,000	,000	,000	,000
foiEmo	,142	,072	,274	,174	,351	,523	,000	,000	,000	,000	,000	,000
foiWert	,167	,085	,321	,204	,412	,613	,000	,000	,000	,000	,000	,000
foiEpi	,142	,072	,273	,174	,351	,522	,000	,000	,000	,000	,000	,000
postEpi	,185	,057	,244	,204	,647	,000	,000	,000	,000	,000	,000	,000
postWert	,217	,067	,287	,241	,761	,000	,000	,000	,000	,000	,000	,000
postEmo	,186	,057	,247	,206	,653	,000	,000	,000	,000	,000	,000	,000
Transparenz	,058	,020	,049	,077	,000	,000	,000	,000	,000	,000	,000	,000
pra20	,084	,618	,281	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra4	,080	,586	,267	,000	,000	,000	,000	,000	,000	,000	,000	,000
Atmosphäre (Einzelitem)	,112	,039	,094	,148	,000	,000	,000	,000	,000	,000	,000	,000
Betreuer (Einzelitem)	,116	,040	,098	,154	,000	,000	,000	,000	,000	,000	,000	,000
pra6	,248	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra12	,132	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra23	,194	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
foa19	,114	,178	,170	,116	,251	,310	,046	,009	,427	,168	,026	,000
foa15	,131	,205	,196	,133	,288	,356	,053	,010	,491	,193	,030	,000
foa1	,107	,167	,160	,109	,235	,290	,044	,008	,401	,157	,025	,000
foa11	,039	,701	,033	,052	,171	,000	,000	,242	,000	,000	,000	,000
foa28	,042	,752	,035	,056	,183	,000	,260	,000	,000	,000	,000	,000
foa23	,031	,559	,026	,041	,136	,000	,193	,000	,000	,000	,000	,000
poa45	,042	,770	,036	,056	,184	,000	,000	,000	,000	,000	,000	,000
poa41	,041	,739	,034	,054	,177	,000	,000	,000	,000	,000	,000	,000
poa26	,031	,564	,026	,041	,135	,000	,000	,000	,000	,000	,000	,000
pra21	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra14	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra1	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
foiPlan	,499	,010	,024	,038	,124	,000	,000	,209	,000	,000	,000	,000
foiDurch	,502	,010	,024	,038	,125	,000	,000	,211	,000	,000	,000	,000
foiAusw	,647	,013	,031	,049	,161	,000	,000	,272	,000	,000	,000	,000
posPlan	,310	,017	,040	,064	,209	,000	,000	,000	,000	,000	,000	,000
posDurch	,387	,021	,050	,079	,261	,000	,000	,000	,000	,000	,000	,000
posAusw	,459	,025	,060	,094	,310	,000	,000	,000	,000	,000	,000	,000
prePlan	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
preDurch	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
preAusw	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
Authentizität	,154	,053	,130	,204	,000	,000	,000	,000	,000	,000	,000	,000
Zusammenarb.	,138	,048	,116	,182	,000	,000	,000	,000	,000	,000	,000	,000
Alltagsbezug	,157	,054	,132	,207	,000	,000	,000	,000	,000	,000	,000	,000

C.3 Modell affektiver Einflussgrößen am Lernort BeLL-Bio – Totale Effekte

Standardisierte Totale Effekte (Strukturgleichungsmodell BeLL Bio; bei den Variablen handelt es sich – wenn nicht explizit Einzelitems genannt sind – um Summenscores)

	preSelbstExp	preFSK	Sachinteresse	Fachinteresse	BELLBio	postAktInteresse	postFSK	postSelbstExp	foiAktInteresse	foiFSK	foiSelbstExp	Berufsinteresse
Sachinteresse	,299	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
Fachinteresse	,090	,658	,000	,299	,000	,000	,000	,000	,000	,000	,000	,000
BELLBio	,230	,079	,193	,304	,000	,000	,000	,000	,000	,000	,000	,000
postAktInteresse	,256	,079	,339	,284	,899	,000	,000	,000	,000	,000	,000	,000
postFSK	,045	,818	,038	,059	,196	,000	,000	,000	,000	,000	,000	,000
postSelbstExp	,647	,035	,084	,133	,437	,000	,000	,000	,000	,000	,000	,000
foiAktInteresse	,197	,100	,379	,241	,487	,724	,000	,000	,000	,000	,000	,000
foiFSK	,045	,802	,038	,059	,195	,000	,277	,000	,000	,000	,000	,000
foiSelbstExp	,786	,016	,038	,059	,196	,000	,000	,000	,000	,000	,000	,000
Berufsinteresse	,143	,224	,214	,146	,315	,389	,058	,011	,537	,211	,033	,000
foiEmo	,142	,072	,274	,174	,351	,523	,000	,000	,722	,000	,000	,000
foiWert	,167	,085	,321	,204	,412	,613	,000	,000	,846	,000	,000	,000
foiEpi	,142	,072	,273	,174	,351	,522	,000	,000	,721	,000	,000	,000
postEpi	,185	,057	,244	,204	,647	,720	,000	,000	,000	,000	,000	,000
postWert	,217	,067	,287	,241	,761	,847	,000	,000	,000	,000	,000	,000
postEmo	,186	,057	,247	,206	,653	,727	,000	,000	,000	,000	,000	,000
Transparenz	,058	,020	,049	,077	,253	,000	,000	,000	,000	,000	,000	,000
pra20	,084	,618	,281	,939	,000	,000	,000	,000	,000	,000	,000	,000
pra4	,080	,586	,267	,891	,000	,000	,000	,000	,000	,000	,000	,000
Atmosphäre (Einzelitem)	,112	,039	,094	,148	,486	,000	,000	,000	,000	,000	,000	,000
Betreuer (Einzelitem)	,116	,040	,098	,154	,506	,000	,000	,000	,000	,000	,000	,000
pra6	,248	,000	,829	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra12	,132	,000	,442	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra23	,194	,000	,648	,000	,000	,000	,000	,000	,000	,000	,000	,000
foa19	,114	,178	,170	,116	,251	,310	,046	,009	,427	,168	,026	,796
foa15	,131	,205	,196	,133	,288	,356	,053	,010	,491	,193	,030	,916
foa1	,107	,167	,160	,109	,235	,290	,044	,008	,401	,157	,025	,747
foa11	,039	,701	,033	,052	,171	,000	,242	,000	,000	,873	,000	,000
foa28	,042	,752	,035	,056	,183	,000	,260	,000	,000	,937	,000	,000
foa23	,031	,559	,026	,041	,136	,000	,193	,000	,000	,696	,000	,000
poa45	,042	,770	,036	,056	,184	,000	,941	,000	,000	,000	,000	,000
poa41	,041	,739	,034	,054	,177	,000	,904	,000	,000	,000	,000	,000
poa26	,031	,564	,026	,041	,135	,000	,690	,000	,000	,000	,000	,000
pra21	,000	,931	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra14	,000	,893	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
pra1	,000	,579	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
foiPlan	,499	,010	,024	,038	,124	,000	,000	,209	,000	,000	,634	,000
foiDurch	,502	,010	,024	,038	,125	,000	,000	,211	,000	,000	,639	,000
foiAusw	,647	,013	,031	,049	,161	,000	,000	,272	,000	,000	,823	,000
posPlan	,310	,017	,040	,064	,209	,000	,000	,479	,000	,000	,000	,000
posDurch	,387	,021	,050	,079	,261	,000	,000	,598	,000	,000	,000	,000
posAusw	,459	,025	,060	,094	,310	,000	,000	,709	,000	,000	,000	,000
prePlan	,644	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
preDurch	,520	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
preAusw	,802	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
Authentizität	,154	,053	,130	,204	,671	,000	,000	,000	,000	,000	,000	,000
Zusammenarb.	,138	,048	,116	,182	,599	,000	,000	,000	,000	,000	,000	,000
Alltagsbezug	,157	,054	,132	,207	,681	,000	,000	,000	,000	,000	,000	,000

Erklärung

Hiermit erkläre ich, diese Arbeit selbständig angefertigt und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet zu haben.

Ich erkläre ebenso, dass diese Dissertation keiner anderen Fakultät mit dem Ziel vorgelegt worden ist, einen akademischen Titel zu erwerben.

Karsten Damerau

(Wuppertal, im April 2012)