

Bedeutung des Protokollierens für den naturwissenschaftlichen Erkenntnisprozess

von

Lisa Engl

geb. am 02.01.1987 in Pirmasens

Angenommene Dissertation zur Erlangung des akademischen Grades eines
Doktors der Philosophie
Fachbereich 7: Natur- und Umweltwissenschaften
Universität Koblenz-Landau

Berichterstatter:

Prof. Dr. Björn Risch, Landau

Prof. Dr. Jürgen Roth, Landau

Tag der Disputation: 03. Februar 2017

*So eine Arbeit wird nie fertig.
Man muss sie für fertig erklären,
wenn man nach Zeit und Umständen
das Möglichste getan hat.*

J. W. Goethe

Inhaltsverzeichnis

1	AUSGANGSLAGE UND PROBLEMSTELLUNG	1
2	DAS PROTOKOLL – DEFINITION, TYPEN UND ANWENDUNGSBEREICHE	4
2.1	PROTOKOLLE IM KONTEXT „SCHULE“	6
2.2	PROTOKOLLE IM KONTEXT „SCHÜLERLABORE“	19
3	DAS PROTOKOLL ALS ELEMENT DES NATURWISSENSCHAFTLICHEN ERKENNTNISPROZESSES	23
3.1	THEORETISCHE ÜBERLEGUNGEN – BILDUNGSSTANDARDS UND SCHREIBEN	23
3.2	EMPIRISCHE STUDIE ZUM SCHREIBEN VON PROTOKOLLEN – DARSTELLUNGSART UND EINFLUSSFAKTOREN	32
3.2.1	METHODEN DER DATENERHEBUNG	38
3.2.2	KONZEPTION DER LERNEINHEIT	41
3.2.3	METHODEN DER DATENAUSWERTUNG	47
3.2.4	UNTERSUCHUNGSERGEBNISSE	52
3.2.5	DISKUSSION DER ERGEBNISSE	63
3.2.6	METHODENKRITIK UND ÄNDERUNGSVORSCHLÄGE	69
3.3	EMPIRISCHE STUDIE ZUR BEDEUTUNG VON PROTOKOLLEN – LERNWIRKSAMKEIT UND QUALITÄT	72
3.3.1	PILOTIERUNG	77
3.3.2	METHODEN DER DATENERHEBUNG	80
3.3.3	KONZEPTION DER LERNEINHEIT	85
3.3.4	METHODEN DER DATENAUSWERTUNG	93
3.3.5	GÜTEKRITERIEN DES KATEGORIENSYSTEMS	102
3.3.6	UNTERSUCHUNGSERGEBNISSE	107
3.3.7	DISKUSSION DER ERGEBNISSE	123
4	ZUSAMMENFASSUNG UND AUSBLICK	129
	LITERATUR	I
	DANKSAGUNG	X
	ABBILDUNGSVERZEICHNIS	XI
	TABELLENVERZEICHNIS	XIII

1 Ausgangslage und Problemstellung

Das Schreiben von Protokollen ist zumeist keine beliebte Aufgabe (SÜSELBECK, 2003). Auf Dienstbesprechungen und Konferenzen verstecken sich alle Teilnehmerinnen und Teilnehmer¹ regelrecht, wenn es darum geht, wer das heutige Protokoll schreibt (ebd.). Auch in der Schule ist der Begriff „Protokoll“ über alle Jahrgangsstufen hinweg negativ belegt (KRAUS & STEHLIK, 2008). Zusätzlich wirft das Protokollieren noch einige didaktische Probleme auf: Ein hoher Zeitaufwand und die mangelnde Beliebtheit macht das Protokollieren Schülern, aber auch Lehrern regelrecht zur Last (ebd.). Jedoch gehört das Protokoll zu der Textart, auf die im naturwissenschaftlichen Unterricht am wenigsten verzichtet werden kann: Weitervermittlung von Wissen als Grundlage zur Nacharbeit oder Klausurvorbereitung oder als Hilfsmittel zur Systematisierung von Beobachtungen naturwissenschaftlicher Phänomene (ebd.). Im schulischen Kontext wird das Protokollieren von Versuchsergebnissen im Unterrichtsgang aufgegriffen, so dass ähnliche Produkte entstehen und die Tätigkeit für alle unter gleichen Bedingungen durchgeführt wird. Beim eigenständigen Experimentieren zum Beispiel im Schülerlabor ist ein gemeinsames Vorgehen beim Protokollieren kaum möglich. Ob die Schüler dadurch überhaupt ihre Ergebnisse in irgendeiner Weise festhalten steht an dieser Stelle zur Diskussion. Entscheidend für die Wirksamkeit von Protokollen scheint deshalb die Einstellung der Schüler zum Schreiben und deren Einsatz während der Experimentierphase zu sein (NIESWANDT, 1998). Denn fast alle Schüler experimentieren gerne, lehnen allerdings die damit verbundenen Tätigkeiten rund um das Protokollieren und Interpretieren von Beobachtungen deutlich ab (FLINTJER, 1998; WITTECK & EIKS, 2004). Die widersprüchliche Einstellung zwischen Funktionalität und Ablehnung von Protokollen macht dieses Thema zu einem interessanten Forschungsgegenstand. Es gilt die Gründe für diesen Antagonismus aufzuklären und bestmöglich zu überwinden.

¹ In dieser Arbeit wird zu Gunsten der Übersichtlichkeit und des Leseflusses lediglich die männliche Form verwendet. Trotzdem wird auf die Gleichstellung beider Geschlechter ausdrücklich hingewiesen.

PROLAB–Protokollieren im Labor

Initialzündung dieses Forschungsthemas war die vom Land Rheinland-Pfalz geförderte Forschungsinitiative II an der Universität Koblenz-Landau im Bereich der Bildungswissenschaften. In der Förderperiode 2012-2013 wurden Lernprozesse unter den Schwerpunkten Protokollieren und Darstellen sowie Erfassen und Identifizieren evaluiert. Als ein wesentliches Element zur Förderung und Entwicklung wissenschaftlichen Denkens und Arbeitens wurde die Fähigkeit identifiziert, Arbeitsprozesse geeignet zu reflektieren und deren Ergebnisse mit Hilfe von passgenauen Repräsentationen so festzuhalten, dass die entstehenden Dokumente sich gewinnbringend für weitere Lernprozesse nutzen lassen. Eine Analyse fachspezifischer Wissensstrukturen und die hinter diesen Wissensstrukturen stehenden Erkenntnismethoden aus der Perspektive der Fachdidaktik, Pädagogik und Psychologie sollte deutlich machen, dass bei allen inhaltlichen Unterschieden auch gemeinsame fachübergreifende Strukturen identifiziert werden können. Seitens der naturwissenschaftlichen Fachdidaktiker wurde der Schwerpunkt des Projektes auf das Protokollieren gelegt und aus unterschiedlichen Perspektiven interdisziplinär erforscht.

Die Ziele dieser Arbeit sind:

1. Der Begriff des Protokolls wird definiert und für den naturwissenschaftlichen Bildungsbereich abgegrenzt. Zusätzlich sollen Anwendungsgebiete von Protokollen für den innerschulischen und außerschulischen Kontext aufgezeigt werden.
2. Mögliche Schwierigkeiten und Unterscheidungsmerkmale bei der Tätigkeit des Schreibens von Protokollen der Schüler werden identifiziert.
3. Die Bedeutung des Protokollierens für den naturwissenschaftlichen Erkenntnisprozess und für die Lernwirksamkeit wird aufgezeigt.
4. Die Qualität von Protokollen wird normativ bestimmt.

Diese Ziele wurden durch folgende Maßnahmen erreicht:

Zur Erarbeitung der Ziele wurde das Projekt „**PROLAB–Protokollieren im Labor**“ initiiert. Dieses Projekt beinhaltete drei Teilstudien, die an außerschulischen Lernorten durchgeführt wurden. Im Unterschied zum Vorgehen in der Schule, erfolgt das Lernen am außerschulischen Lernort lediglich durch bereitgestellte Materialien und ohne anleitende Lehrperson. Während Teilstudie 1 und 2 einen explorativen Charakter aufweisen, liegt die Hauptaussagekraft dieser Dissertation Teilstudie 3 zu Grunde.

Teilstudie 1 Eine Übersicht an Definitionen und Anwendungsgebieten von Protokollen im innerschulischen Bereich der naturwissenschaftlichen Fachdidaktik wird durch die Literatur aufgezeigt. Für den Bereich des außerschulischen Einsatzes von Protokollen erfolgt eine Erhebung des Status Quo zum Umgang mit Protokollen „Protokolle im Kontext Schülerlabor“.

Teilstudie 2 Im Rahmen des Forscherverbundes der Forschungsinitiative II entstanden erste Kenntnisse über das Protokoll und die Tätigkeit des Protokollierens von Schülern. Die explorative Herangehensweise an das „Schreiben von Protokollen“ ermöglichte Unterschiede in den Protokollen zu identifizieren und durch gegebene Strukturen der Untersuchung zu erläutern. Zusätzlich sollte das Beobachten und Befragen der Schüler deren Blickwinkel auf das Protokollieren und die damit verbundenen Schwierigkeiten ermöglichen. Um Schülern beim Verschriftlichen der Ergebnisse eines Experiments Hilfestellungen zu geben, wurden Unterstützungsmaßnahmen bei der unbeliebten Tätigkeit des Protokollierens eingesetzt.

Teilstudie 3 Das Protokoll ist ein Teil des naturwissenschaftlichen Erkenntnisprozesses. Um jedoch diesen Teilschritt einer Notwendigkeit für die Lernwirksamkeit zuschreiben zu können, wurde mit Teilstudie 3 die „Bedeutung von Protokollen“ überprüft. Deshalb soll an dieser Stelle die Auswirkung eines fehlenden Protokolls und qualitativ unterschiedlicher Protokolle auf den Lernerfolg untersucht werden. Zur Bestimmung der Qualität von Protokollen wurde in dieser Studie ein Messinstrument entwickelt und anschließend auf die Gütekriterien Objektivität, Reliabilität und Validität überprüft.

In dieser Arbeit werden unterschiedliche Definitionen, Typen und Anwendungsgebiete von Protokollen im Kontext „Schule“ (Kapitel 2.1) vorgestellt. Im Anschluss wird die Verwendung von Protokollen im Kontext „Schülerlabore“ (Kapitel 2.2) durch Teilstudie 1 empirisch erörtert. Kapitel 3 zeigt das Protokoll als Element des naturwissenschaftlichen Erkenntnisprozesses auf. Anhand der Bildungsstandards und des Schreibens im naturwissenschaftlichen Unterricht wird die Bedeutung des Protokolls theoretisch (Kapitel 3.1) erläutert. Anschließend werden zwei weitere empirische Studien vorgestellt. Teilstudie 2 (Kapitel 3.2) untersucht das Schreiben von Protokollen durch die Darstellungsart und die Einflussfaktoren des Protokollierens. Die Bedeutung von Protokollen für die Lernwirksamkeit und die Bedeutung der Qualität von Protokollen wird anhand von Teilstudie 3 (Kapitel 3.3) dargestellt. Kapitel 4 fasst die Vorgehensweisen und Ergebnisse der Arbeit zusammen und gibt einen Ausblick auf weitere Untersuchungsmöglichkeiten.

2 Das Protokoll – Definition, Typen und Anwendungsbereiche

Das Protokoll ist ein Begriff der über die Naturwissenschaften hinaus geprägt ist und nimmt dennoch im naturwissenschaftlichen Bereich eine wichtige Stellung ein. Das Wort Protokoll ist im deutschen Standardwortschatz seit dem 16. Jahrhundert nachweisbar und dem mittellateinischen *protocollum* angelehnt, welches wiederum selbst dem mittelgriechischen *protókollon* (aus *prótos* „erster“ und *kólla*, „Klebe, Leim“), der Ursprungsbedeutung „[den amtlichen Papyrusrollen] vorgeleimtes Blatt“ entspricht (KLUGE, 2011). Laut BIBLIOGRAPHISCHEN INSTITUT (2013) legt ein Protokoll fest, zu welchem Zeitpunkt oder in welcher Reihenfolge ein Vorgang durchgeführt wird. Die Niederlegung dieser Tätigkeit erfolgt in Schriftform (ebd.). Besonders in der Fachsprache wird das Protokoll als genauer Bericht über den Verlauf und die Ergebnisse eines Versuchs oder einer Operation definiert (ebd.). Im Bildungsbereich existieren verschiedene Begriffe über die Möglichkeit durch Protokolle etwas aufzuschreiben und alle haben ihre eigene Bedeutung sowie ihre spezifische Einsatzmöglichkeit. Eine ausführliche Darstellung über die Definition und das Anlegen von Protokollen wird vor allem in der Fachdidaktik Deutsch dargestellt (BRÜNING, 1990). MOLL (2003) gibt einen Überblick über verschiedene Protokolltypen im nicht naturwissenschaftlichen Bereich. Dazu zählen zum Beispiel Verbalprotokolle, Sitzungsprotokolle, Ergebnisprotokolle oder Gedächtnisprotokolle. Je nach Anlass und Nutzen des Protokolls kann ein bestimmter Protokolltyp gewählt werden. Allgemein zählt das Protokoll in der Schule als Werkzeug zur Kommunikation zwischen Lernenden oder auch zwischen Lehrenden und Lernenden. Das Protokollieren verknüpft die Kompetenzbereiche der Erkenntnisgewinnung und Kommunikation. Beim Protokollieren steht im Fokus, das Wesentliche der Teilschritte des Erkenntnisprozesses zu erfassen, sowie geeignet festzuhalten und darzustellen. Protokolle haben das Ziel, möglichst knapp aber genau den Weg der naturwissenschaftlichen Erkenntnisgewinnung zu belegen (KRAUS & STEHLIK, 2008). Doch genau diese präzise Art der Dokumentation stellt Schüler jedoch vor schwierige Herausforderungen (MOLL, 2003; STEINHOFF, 2003; PRECHTL, 2008).

Die Protokollinhalte können durch verschiedene Repräsentationsformen dargestellt werden. Der Umgang mit Repräsentationsformen sowie der Wechsel zwischen Repräsentationsformen sind deshalb für die Arbeit mit Protokollen zu berücksichtigen. Repräsentationen sind Abbilder eines Gegenstandes oder einer Handlung. Unterschiedliche Repräsentationen der gleichen Handlung können unterschiedliche Informationen beinhalten (SCHNOTZ, BAADTE, MÜL-

LER & RASCH, 2011). Nach SCHNOTZ (2005) lassen sich depiktionale (bildliche) und deskriptionale (beschreibende) Repräsentationen unterscheiden. Depiktionale Darstellungsformen sind zum Beispiel realistische Bilder oder schematische Zeichnungen. Deskriptionale oder auch symbolische Repräsentationen bestehen aus Zeichen, die keine Ähnlichkeit mit dem zu repräsentierenden Gegenstand aufweisen. Dazu gehören aus Symbolen bestehende Repräsentationen wie Text, Tabellen oder Formeln. Hier seien auch mathematische oder chemische Symbole genannt (vgl. KOZMA & RUSSELL, 2005a). Das Nutzen von verschiedenen Repräsentationsformen als Werkzeug und der Wechsel zwischen den Darstellungsformen ist im naturwissenschaftlichen Denken elementar (SCHEID, 2013). Die Repräsentationskompetenz beinhaltet zum einen die Fähigkeit mit vorgegebenen Darstellungen umzugehen und zum anderen die Fähigkeit zum selbstständigen Erzeugen von Darstellungen zu einem Lerninhalt (COX, 1999; KOZMA & RUSSELL, 2005b; NEUBERT, 2011; SCHNOTZ ET AL., 2011). Mit dieser Fähigkeit eines Lernalters können Darstellungsformen erzeugt, gewechselt und verbunden werden (SCHEID, 2013). Das Verstehen von Experimenten und Phänomenen hängt eng mit dieser Fähigkeit zusammen (ebd.). Ein Hauptteil der genannten Autoren bezieht sich auf den Umgang mit Repräsentationen und dem Wechsel zwischen unterschiedlichen Repräsentationsformen (Repräsentationaler Kohärenz). In dieser Arbeit werden Repräsentationsformen als Werkzeug zum Erstellen von Protokollen angesehen. Eine Untersuchung und Bewertung der repräsentationalen Kohärenz ist deshalb nicht notwendig.

2.1 Protokolle im Kontext „Schule“

In der mathematisch-naturwissenschaftlichen Fachdidaktik und in Schulbüchern existieren unterschiedliche Begrifflichkeiten für ein Protokoll: Versuchsprotokoll, Verlaufsprotokoll, Leitfadenprotokoll, Lernprotokoll, Phänomenprotokoll und Handlungsprotokoll. Tabelle 1 stellt einen Überblick über diese Protokollformen, deren Beschreibung, Einsatzmöglichkeiten und die relevante Literaturverweise im Überblick dar.

Tabelle 1: Definition der verschiedenen Protokollformen, deren Einsatzmöglichkeiten und die verwendete Literatur.

Protokollform	Beschreibung des Inhalts	Einsatzmöglichkeiten	Literatur
Versuchsprotokoll	Ergebnisse und Verlauf eines Experimentes	Planung, Durchführung & Deutung von Experimenten	BRÜNING (1990)
Verlaufsprotokoll	Beschreibung des experimentellen Vorgehens	Prozessorientierte Leistungsbewertung	EMDEN & SUMFLETH (2012)
Leitfadenprotokoll	Leitfragen gestützte Erarbeitung des Experimentes	Erlernen der Protokollier-techniken	KRAUS & STEHLIK (2008)
Lernprotokoll	Reflexion über den eigenen Lernprozess	Ausdrucks- und Reflexionsmittel von Lernprozessen	BERTHOLD, NÜCKLES, RENKL (2004/2007) NÜCKLES, SCHWONKE, BERTHOLD & RENKL (2004)
Phänomenprotokoll	Beschreibung und Bedingungen eines Phänomens	Verknüpfung von Phänomenen und Kompetenzen	REHM (2006/2007)
Handlungsprotokoll	Versprachlichung eines kognitiven Prozesses	Denkwerkzeug zur Begriffsentwicklung	DÖRFLER (2003/2012)

2.1.1 Protokollformen

Nachfolgend werden die sechs verschiedenen Protokollformen mit Beschreibung der Inhalte, Anwendungsbereiche und Forschungsergebnissen zusammenfassend beschrieben.

Versuchsprotokoll

Klassisch wird im naturwissenschaftlichen Unterricht und Studium zunächst an das Versuchsprotokoll gedacht. In einem Versuchsprotokoll wird die Planung und Durchführung eines Experimentes sowie dessen Ergebnisse festgehalten (BRÜNING, 1990). Diese Form des Protokolls ist eine Verbindung zwischen Experiment und Protokoll (vgl. Tabelle 2) und erfolgt nach BRÜNING (1990) strukturgleich. Das Versuchsprotokoll lehnt sich an die Struktur des Experimentes an und hat immer den gleichen Aufbau.

Tabelle 2: Vergleich der Strukturen eines Experimentes und des dazugehörigen Protokollinhaltes.

Experiment	Protokoll
Planung	1. Aufgabe (Problem, Hypothesen, Ziel des Versuchs) 2. Versuchsaufbau
Durchführung	3. Versuchsdurchführung 4. Versuchsbeobachtungen, Versuchsdaten, Messresultate
Analyse	5. Versuchsauswertung (qualitativ, quantitativ) 6. Versuchsergebnis, Fehlerabschätzung, Diskussion

Die Protokollstruktur ist für die Schüler ein wiederkehrendes Element und somit Hilfestellung beim Experimentieren (ebd.). Deshalb gibt es dafür auch zahlreiche Vorlagen und Muster, die den Inhalt und die Form vorgeben. In vielen naturwissenschaftlichen Schulbüchern wird ein Beispiel für ein Protokoll und die notwendigen Teilbereiche abgebildet. Abbildung 1 zeigt ein Muster mit den einzelnen Elementen eines Versuchsprotokolls aus dem Schulbuch PRISMA CHEMIE 7-10 (2006).

Thema des Versuchs
Der Versuch wird als Thema oder Fragestellung formuliert. Es sollte möglichst in einem Satz, aber exakt formuliert werden.

Material
Es werden alle benötigten Chemikalien, Geräte und sonstige Hilfsmittel notiert.

Name und Datum
Wie bei jeder Protokollerstellung enthält auch ein Laborprotokoll den Namen des Protokollanten und das Datum der Anfertigung.

Sicherheitsmaßnahmen
Bei der Durchführung des Versuchs müssen bestimmte Sicherheitsmaßnahmen beachtet werden. Das Tragen einer Schutzbrille ist fast immer unerlässlich. Sind die verwendeten Chemikalien Gefahrstoffe, müssen die entsprechenden R- und S-Sätze vorher durchgelesen und nötige Sicherheitsvorkehrungen getroffen werden.

Beobachtungen
Alle Einzelheiten, die während des Versuchsablaufs beobachtet werden, insbesondere die Messergebnisse werden notiert.

Entsorgung
Häufig bleiben nach Beendigung des Versuchs Abfälle (Chemikalienreste, Gerätebruch) übrig. Diese müssen in geeigneter Form entsorgt werden.

Versuchsaufbau
Besonders geeignet zur Darstellung einer Versuchsskizze ist eine Labor-Geräteschablone oder ein PC-Programm für Laborgeräte. Versuchsskizzen sollten beschriftet werden.

Versuchsanleitung
Die Durchführung des Versuchs wird in kurzer, aber verständlicher Form beschrieben.

Auswertung
Die Auswertung erhält man aufgrund der Versuchsbeobachtungen. Sie besteht häufig aus einer allgemeinen Formulierung oder einer Gesetzmäßigkeit.

Abbildung 1: Vorlage für ein Versuchsprotokoll aus einem Schulbuch. Quelle: PRISMA CHEMIE 7-10, 2006.

Versuchsprotokolle beinhalten die Planung und Auswertung eines Versuchs, jedoch nicht das Vorgehen der Schüler beim Aufbauen und Messen (SCHREIBER & THEYßEN, 2015). Mit dem Konstrukt Versuchsprotokoll sind bereits Vorstellungen verbunden, die regeln, wie das Produkt aussehen soll beziehungsweise nicht aussehen darf (PRECHTL, 2010). KRAUS & STEHLIK (2008) führten eine Untersuchung über die Assoziationen von Schülern zu dem Begriff „Protokoll“ durch. Über vier Jahrgänge hinweg verbanden die Schüler mit dem Begriff den Aufbau eines Versuchsprotokolls (ebd.). Dem Anfertigen von Versuchsprotokollen wird aufgrund der Anwendung in Forschung und Unterricht eine hohe Bedeutung zugesprochen (PRECHTL, 2008). Jedoch sind uns keine Studien darüber bekannt.

Nach der Methode dieser klassischen Definition wird im naturwissenschaftlichen Unterricht am häufigsten ein Protokoll erstellt. Vorteil des Versuchsprotokolls ist die vorgegebene inhaltliche Gliederung, an der sich gut orientiert werden kann. Im Unterricht wird dieser Weg mit den Schülern gemeinsam nachvollzogen und durchgeführt. Die Anwendung dieser Protokollform ist sinnvoll, wenn genau diese Schritte von Fragestellung bis hin zur Erklärung durchlaufen werden. Ist es jedoch vorgesehen ein bestimmtes Experiment nur durchzuführen und zum Beispiel keine Fragestellung oder Erklärung zu besprechen, ist diese Protokollform ungeeignet. Diese Vorgehensweise ist vor allem in Schülerlaboren zu beobachten.

Verlaufsprotokoll

Die Möglichkeit das Vorgehen der Schüler beim Experimentieren zu dokumentieren und damit die Leistung der naturwissenschaftlichen Arbeitsweisen zu bewerten, zeigen EMDEN & SUMFLETH (2012) auf. In einem offenen Protokollformat protokollieren die Schüler während des Experimentierens mehrmals für circa drei Minuten (ebd.). Diese Protokollform entspricht der Klassifizierung aus der Deutsch-Didaktik: In einem Verlaufsprotokoll werden wichtige Tatsachen, Äußerungen und Beschlüsse wiedergegeben (BRÜNING, 1990). Die Protokollbögen sind vorstrukturiert durch die Schritte des experimentellen Arbeitens (Idee finden, Experiment durchführen, Schlussfolgern) und eine Beurteilung über erfolgreiche und nicht erfolgreiche Experimentieransätze (Abbildung 2).

Verlaufsprotokoll

1. Ideen (Welche Möglichkeiten gibt es, das gestellte Problem zu lösen?)

Bsp: 3 min: Wir haben die Idee das Salzwasser durch einen Filter in ein anderes Becherglas zu gießen.

ODER 3 min: Salz aus Wasser filtern

3 min	_____
6 min	_____
9 min	_____
12 min	_____
15 min	_____
18 min	_____
20 min	_____

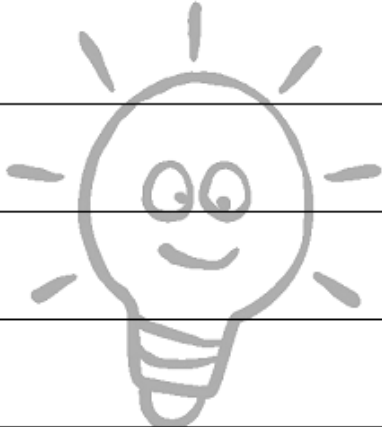


Abbildung 2: Beispiel für ein Verlaufsprotokoll. Quelle: EMDEN & SUMFLETH, 2012.

Ein Vergleich der Experimentierleistungen durch Lernprozessgrafiken aus Videos mit der Bewertung der Verlaufsprotokolle zeigt einen Zusammenhang zwischen den beiden Leistungsmessungen (EMDEN & SUMFLETH, 2012). Die Autoren empfehlen eine Einführung dieser Protokollmethode, da ansonsten eine kognitive Überlastung durch den Umgang mit der unvertrauten Methode möglich ist (ebd.).

Das Verlaufsprotokoll verfolgt die gleiche inhaltliche Struktur wie das Versuchsprotokoll, ist aber auf drei Komponenten reduziert. Zusätzlich wird hier die Idee verfolgt, dass während des ganzen Experimentes die eigene Arbeitsweise dokumentiert und dadurch reflektiert wird. Diese Protokollform orientiert sich an den eigenen Ideen der Schüler und zeigt sich bei der Problemlösung eigener Fragen als zielführend. Jedoch ist sie ebenso ungeeignet für vorgegebene Experimente ohne einen Prozess der Problemlösung und somit für den Gebrauch in Schülerlaboren.

Leitfadenprotokoll

KRAUS & STEHLIK (2008) schlagen aufgrund der mangelnden Funktionalität von Versuchsprotokollen im Schulalltag die Arbeit mit „Leitfadenprotokollen“ vor. Das ist eine Mischform aus wissenschaftlichem Protokoll und Gesprächsprotokoll, welche Leitfragen enthält und dadurch Fehlvorstellungen beim Experimentieren vermeiden soll (ebd.). Abbildung 3 zeigt Kontrollfragen zur Entwicklung eines solchen Protokollformates, welches sich vor allem an der Lernsystematik der Schüler orientiert.

- *Was war das Thema?*
- *Welche Gesichtspunkte/Unterthemen wurden angesprochen?*
- *Welche Thesen wurden vertreten? (Welche Hypothesen entwickelt?)*
- *Wie werden sie begründet? (Argumente)*
- *Wie werden sie veranschaulicht? (Beispiele, Belege, Experimente)*
- *Welche Gesprächsphasen lassen sich unterscheiden?*
- *Zu welchen Ergebnissen/Zwischenergebnissen ist man gekommen?*

Abbildung 3: Kontrollfragen zur Entwicklung von Leitfadenprotokollen. Quelle: KRAUS & STEHLIK, 2008.

Das Leitfadenprotokoll findet Anwendung in der Förderung des Protokollschreibens und Erlernen des wissenschaftlichen Arbeitens. Die Überprüfung dieser Methode zeigte, dass den Schülern einzelne Schritte des Protokollierens deutlich gemacht werden, jedoch mehrere Unterrichtsstunden zu diesem Protokolltraining eingeplant werden sollten (KRAUS & STEHLIK, 2008).

Das Leitfadenprotokoll ist eine offenere Protokollform. Fragen geben Impulse für das Verschriftlichen eines Experimentes oder des Problemlösens. Somit handelt es sich hier um eine Mischform von Versuchs- und Verlaufsprotokoll mit verschiedenen Einsatzmöglichkeiten. Jedoch ist laut der Autoren für das Protokollieren mit dem Leitfadenprotokoll Übung notwendig. Für den einzelnen Einsatz im Unterricht, empirischen Studien und außerschulischen Lernorten ist diese Art des Protokollierens deshalb ungeeignet.

Lernprotokoll

Die Freiburger Psychologen NÜCKLES, SCHWONKE, BERTHOLD & RENKL (2004) verwenden den englischen Begriff „learning protocol“ und untersuchen die Wirkung dieser Protokollart auf den Lernerfolg. Nach NÜCKLES ET AL. (2004) ist ein Lernprotokoll eine schriftliche Reflexion des Lernenden über den eigenen Lernprozess und dessen Ergebnis. Werden Lernprotokolle über einen längeren Zeitraum angefertigt, erfolgt die Bezeichnung „Lerntagebuch“ (ebd.). Das Lernprotokoll soll den Lernenden durch die Strukturierung von Texten auch ihr Wissen neu organisieren (BERTHOLD, NÜCKLES, RENKL, 2004). Im Vergleich zu einer verbalen Wiedergabe über den Lernstoff bleibt dieser erhalten und gibt dadurch die Möglichkeit die Inhalte erneut zu lesen, zu überdenken und sie durch die Anwendung von kognitiver Aktivität in Wissen zu transformieren (*learning by writing*) (BERTHOLD, NÜCKLES & RENKL, 2007). Jedoch zeigt die Arbeit von NÜCKLES ET AL. (2004), dass das Anfertigen eines Lernprotokolls nicht gleichzeitig eine höhere Lernaktivität garantiert. In einer Video-Studie mit Psychologiestudenten wurde gezeigt, dass der Einsatz von Lernprotokollen gemeinsam mit Prompts (Anregungen zum Schreiben) den Lernerfolg und die kognitive Lernaktivität steigern können (BERTHOLD ET AL., 2007).

Bei dieser Protokollform steht der Lernprozess im Vordergrund, weniger das Experiment und dessen Ergebnisse. Deshalb sollte sich für das Protokollieren Zeit genommen werden, um den Prozess reflektieren zu können. Um mehrere Experimente nacheinander durchzuführen, ist der Einsatz des Lernprotokolls zum Beispiel für den Einsatz an außerschulischen Lernorten nicht möglich.

Phänomenprotokoll

REHM (2006 & 2007) arbeitet mit dem Begriff Phänomenprotokoll, welches sich am Konzept der Phänomengeographie orientiert. Dieses Konzept geht davon aus, dass Lernende ein Phänomen unterschiedlich erfahren und dies sich in Form von Handlungen und Äußerungen zeigt (REHM, 2006). Ein Phänomenprotokoll besteht aus der Beschreibung eines Phänomens durch den Beobachter und die Bedingungen unter denen das Phänomen wahrgenommen wurde (REHM, 2007). Der Text des Protokolls wird mit Hilfe von Kompetenzmodellen auf mehrere Qualitätsstufen des Verstehens untersucht (ebd.). Der Autor erhofft sich dadurch die Verknüpfung von Phänomenprotokollen und Kompetenzen, um somit Verstehensprozes-

se zu erfassen. Die Auswertung dieser Protokollform bei Studierenden zeigt, dass eine Einordnung der Fähigkeit in vier Niveaustufen möglich ist (REHM, 2007).

Die Grundidee des Phänomenprotokolls, dass in der Beschreibung von Phänomenen und dem Handeln des Protokollanten sein Verstehen äußert, scheint sinnvoll. Diese Art des Protokollierens wurde jedoch nur mit Studierenden getestet. Deren Verständnis der Phänomene ist wesentlich komplexer und deshalb können unterschiedliche Niveaustufen im Protokoll aufgezeigt werden. Diese Protokollform ist deshalb nicht für alle Altersstufen bei Schülern anwendbar. Außerdem gestaltet sich die Auswertung von Phänomenprotokollen wahrscheinlich als aufwendig. Dies war der Literatur jedoch nicht zu entnehmen.

Handlungsprotokoll

Nach DÖRFLER (2012) sind Handlungsprotokolle der Ausdruck der eigenen Sichtweise und Handlungen des Lernenden. Durch sprachliche Beschreibungen, Grafiken, Symbole und weiteres, werden die Handlungen erfasst, reflektiert und neu organisiert (DÖRFLER, 2012). Dieser kognitive Prozess besteht aus dem Erstellen von Notizen, Aufzeichnungen und Beschreibungen, um die Schritte der Durchführung, Bedingungen der Handlung, Ergebnisse und Produkte zu erfassen (DÖRFLER, 2003). Schüler entwickeln mit Hilfe ihrer Versprachlichung und Konstruktion der eigenen Lernprozesse mathematische Begriffe (ebd.). Die Notwendigkeit von Handlungsprotokollen wird ausschließlich auf theoretischer Grundlage diskutiert und basiert nicht auf empirischen Ergebnissen.

Das Handlungsprotokoll bietet die Möglichkeit Darstellungsform und Inhalte frei wählen zu können. Der Protokollant bestimmt die für ihn wichtigen Begriffe und übt sich dadurch in Selbstorganisation und Reflektion seiner Handlung. Diese Art der Verschriftlichung stellt somit die freieste Form der Protokolle da. Das Einüben dieser Technik scheint nicht notwendig. Nachteil des Handlungsprotokolls für eine Einbindung in den naturwissenschaftlichen Bereich ist, dass sich die Inhalte an einem Experiment orientieren müssen und nicht der freien Konstruktion von Begriffen dienen. Deshalb sollte sich eine Bewertung der Inhalte auf Korrektheit anschließen.

Eine neue Protokollform – das Erarbeitungsprotokoll

Für den Einsatz von Protokollen in der Schule und an außerschulischen Lernorten sollte zunächst die Intention des Einsatzes bedacht werden. Im Landauer Schülerlabor „Nawi-Werkstatt“ liegt der Fokus auf dem selbstständigen Durchführen der Experimente durch die Schüler. Es wird nicht schrittweise ein Erkenntnisweg verfolgt, weshalb Versuchs- und Verlaufsprotokoll sich als nicht sinnvoll herausstellen. Zeitliche Aspekte und Intention des Schülerlabors schließen das Lernprotokoll und das Leitfadenprotokoll aus. Das Phänomenprotokoll ist für die Zielgruppe der 5. und 6. Klassenstufe ungeeignet. Um den Schülern die Darstellungsart und den Inhalt der Protokolle freistellen zu können, müsste die Form der Handlungsprotokolle auf eine experimentelle Arbeitsweise angepasst werden. Eine neue Protokollform wurde deshalb für den Einsatz im Schülerlabor entwickelt: Das Erarbeitungsprotokoll (siehe Kapitel 3.2).

2.1.2 Hilfestellungen und Erlernen von Protokolliertechniken

Protokolle im Unterricht sind für Lehrer und Schüler mit hohem Zeitaufwand und mangelnder Beliebtheit belegt, aber aufgrund ihrer Dokumentationsfunktion unerlässlich (KRAUS & STEHLIK, 2008). Nun stehen alle naturwissenschaftlichen Fächer vor dem gleichen Problem: Wie können Schüler für das Protokollieren motivieren werden? Es gibt wenige Tätigkeiten, die mit so vielen Vorschriften versehen sind, wie das Protokollieren. Deshalb gilt es dem Protokollanten in verschiedenen Bereichen durch Vorlagen oder durch spielerische Vorgehensweise Hilfestellungen zu geben und dadurch die Einstellung zum Protokollieren zu verbessern. In der Naturwissenschaftsdidaktik gibt es in den letzten Jahren diverse Forschungsprojekte, mit dem Ziel das Protokollieren attraktiver zu gestalten. Diese Unterstützungsmaßnahmen sind jeweils abhängig von der Definition des Protokollbegriffs. Tabelle 3 gibt einen Überblick über die Intentionen der unterschiedlichen Hilfestellungen abhängig von ihrer Einsatzmöglichkeit je nach Protokolldefinition.

Tabelle 3: Mögliche Unterstützungsmaßnahmen beim Protokollieren

Hilfestellung	Intention	Protokollform	Autor
Foto-Story	Sprachliche Erleichterung	Versuchsprotokoll	PRECHTL (2008)
1-2-4-Alle-Methode	Sprachliche Erleichterung	Versuchsprotokoll	WITTECK & EIKS (2004)
Satzbausteine, Mustertexte	Sprachliche Erleichterung, Schreibtraining	Versuchsprotokoll	KRÄMER (2011)
Satzbausteine, Lückentexte	Sprachliche Erleichterung, Dauerhafte Hilfestellung	Versuchsprotokoll	HOPPE, KRÄMER & REH (2013)
Prompts	Strukturierungshilfe beim Schreiben	Versuchsprotokoll	MARSCHNER, THILLMANN, WIRTH & LEUTNER (2012) WALPUSKI & SUMFLETH (2007)
		Lernprotokoll	BERTHOLD ET AL. (2007)
Reflexion und Überarbeitung	Schreibtraining	Leitfadenprotokoll	(KRAUS & STEHLIK, 2008)

Der Ansatz von PRECHTL (2008) Versuchsprotokolle als Bildergeschichten zu gestalten, hilft den Schülern auch ohne viele Worte ihre Ergebnisse festzuhalten. Die Foto-Stories enthalten typische Comic-Elemente, wie zum Beispiel Sprech- und Denkblasen, sollen aber im Vergleich zu den fiktionalen Comics eine reale Erfahrung widerspiegeln (ebd.). Abbildung 4 zeigt einen Leitfaden zu dieser Unterstützungsmethode, der unterschiedliche Lerntypen ansprechen soll und eine exemplarisches Ergebnis. Diese Methode soll jedoch nicht einmalig eingesetzt werden, da eine gewisse Übung im Umgang mit Bildergeschichten vorausgesetzt wird (PRECHTL, 2008).

Zeichne das Experiment, das du gerade durchgeführt hast, als „Physik-Foto-Story“ (Bildergeschichte). In dieser Foto-Story spielst du mit. Auch deine Partnerin bzw. dein Partner ist dabei. Du erzählst dem Leser, was ihr gemacht und erlebt habt. Beim Zeichnen solltest du diese zwei Regeln beachten:

- Schreibe in Sprechblasen hinein, worüber ihr während des Experiments gesprochen habt.
- Notiere unter die Bilder einen kurzen Kommentar oder wichtige Hinweise zu deinem Bild.

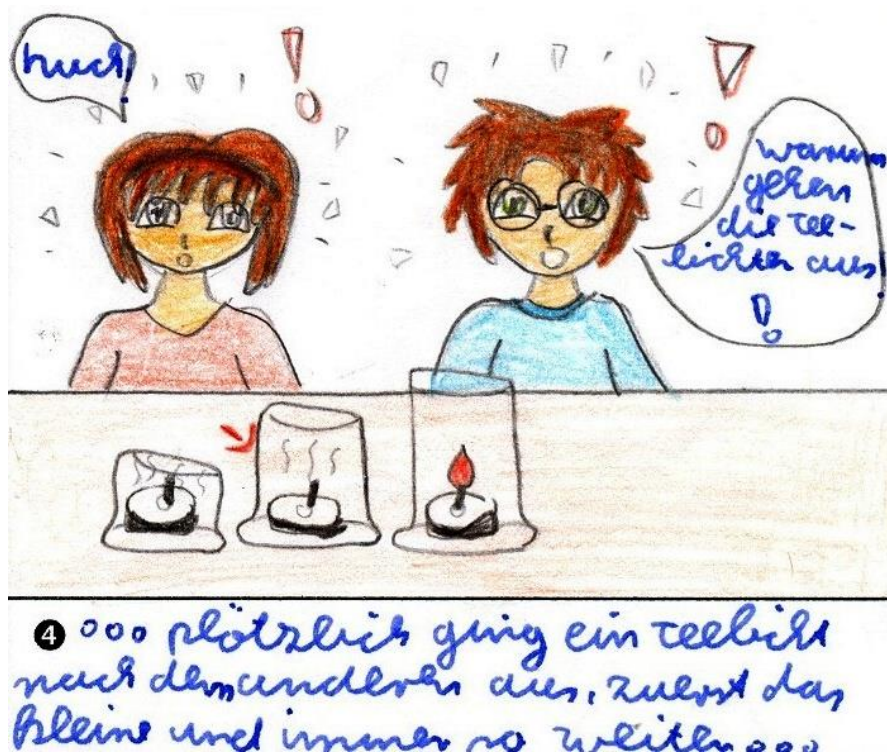


Abbildung 4: Anleitung zur Gestaltung einer Foto-Story. Quelle: Prechtl, 2008.

Um Probleme bei der Formulierung während des Protokollierens zu vermeiden, schlagen WITTECK & EIKS (2004) vor, Versuchsprotokolle in Gruppenarbeit durch die 1-2-4-Alle-Methode zu erstellen. Jeder Schüler schreibt zunächst sein eigenes Versuchsprotokoll, welches der jeweilige Sitznachbar anschließend korrigiert (ebd.). Mit einem zweiten Schülerpaar werden die Protokolle ein weiteres Mal besprochen und ein auf Folie kopiertes Protokoll durch Zerschneiden neu geordnet (ebd.). Die Lösung der Vierer-Gruppe wird auf dem Overhead-Projektor der Klasse präsentiert und mit anderen Lösungsansätzen diskutiert. Bei mehrfacher Durchführung dieser Methode konnten die Autoren Fortschritte beim Erstellen von Versuchsprotokollen beobachten, jedoch nicht empirisch belegen.

Eine weitere Möglichkeit die Schüler beim Protokollieren zu unterstützen, ist die Scaffolding-Methode. Bei Scaffolding-Methoden werden Denkanstöße oder Anregungen zur Erstellung

eines Textes geliefert, welche jedoch unterschiedlich gestaltet sind. Nach diesem Konzept erlernen Schüler mit einem systematischen Training nach KRÄMER (2011) im Verlauf ihrer Schullaufbahn das Schreiben von Protokollen. Mustertexte und Satzteile dienen zur Orientierung beim Schreibtraining der Protokolle, welche im Anschluss in einer Schreibkonferenz überarbeitet werden (ebd.). Durch diese sprachbildenden Maßnahmen erwartet KRÄMER (2011) die Entwicklung einer Protokollschreibkompetenz.

Für die Einführung des Protokollierens in der Orientierungsstufe setzen HOPPE, KRÄMER & REH (2013) schrittweise Hilfestellungen ein. Die Hilfsmittel enthalten zum Beispiel Satzbausteine, den Aufbau eines Versuchsprotokolls oder auch Lückentexte in verschiedenen Schwierigkeitsstufen (vgl. Abbildung 5). Um die individuellen Fähigkeiten kontinuierlich zu fördern, entwickeln die Schüler ein Lernplakat zum Schreiben eines Versuchsprotokolls, welches im Klassenraum als dauerhafte Hilfestellung dienen soll (HOPPE ET AL., 2013).

Hilfeblatt
Welche Teile hat ein Versuchsprotokoll?





Symbol	Protokollteil	Sprachbausteine
	Das Problem	Frage
	Die Vermutung	Ich vermute, dass ...
	Die Materialien und Geräte	Aufzählung ohne Artikel
	Der Versuchsaufbau	Beschriftung mit Linien

Abbildung 5: Hilfeblatt zur Erstellung eines Versuchsprotokolls. Quelle: HOPPE ET AL., 2013.

Die Studien von MARSCHNER, THILLMANN, WIRTH & LEUTNER (2012) oder WALPUSKI & SUMFLETH (2007) zeigen, dass Prompts beziehungsweise Strukturierungshilfen den Schüler beim Experimentieren oder Lernen unterstützen. Der Einsatz von Prompts (Anregungen) in Form von Fragen oder Tipps kann während des Experimentierens Wissen aktivieren, das beim Lerner vorhanden ist, spontan aber nicht genutzt werden (MARSCHNER ET AL., 2012). Deshalb setzen BERTHOLD ET AL. (2007) Prompts auch für das Erstellen von Lernprotokollen ein. Durch die lernstrategischen Fragen werden die Schüler angehalten ihre Lerninhalte zu strukturieren

und gegebenenfalls zu überarbeiten (ebd.). Prompts können durch die Untersuchungen der Autoren als geeignete Methode zur Unterstützung der Lernenden beim Schreiben von Lernprotokollen postuliert werden.

Der Einsatz eines Leitfragenprotokolls im Allgemeinen stellt nach KRAUS & STEHLIK (2008) eine Hilfestellung zum Protokollschreiben dar. Durch die Überarbeitung des Leitfadenprotokolls in Reflexionsphasen findet eine Anknüpfung an die Lernsystematik und ein dauerhaftes Training des Schreibens statt (ebd.). Eine Möglichkeit die Methode zu verinnerlichen, könnte ein geplantes vierstündiges Protokolltraining sein.

Diese Hilfestellungen entstanden infolge eines gemeinsamen Problems: Das Protokollieren ist eine unbeliebte und anstrengende Tätigkeit, die den Schülern erleichtert werden muss. Daher soll mit Hilfe dieser Techniken die Anstrengungsbereitschaft der Schüler verbessert werden. Die unterschiedlichen Hilfsmittel wurden vor allem für den Einsatz mit Versuchsprotokollen konstruiert. Dies veranschaulicht den häufigen Einsatz des Versuchsprotokolls in der Schule und die daraus resultierende notwendige Unterstützungen für Schüler und Lehrer. Ebenso wird deutlich, dass die meisten Hilfestellungen auf eine sprachliche Erleichterung abzielen. Dadurch wird es den Schülern ermöglicht, sich auf die Inhalte zu konzentrieren. Die Idee des Leitfadenprotokolls ermöglicht das schrittweise Erstellen des Protokolls. Jedoch müssen alle Hilfstechniken trainiert und geübt werden. Der einmalige Einsatz wird von keinem der Autoren diskutiert.

2.2 Protokolle im Kontext „Schülerlabore“

Außerschulischen Lernorten, wie beispielsweise Schülerlabore, wurden in den letzten zehn Jahren im naturwissenschaftlichen Unterricht eine große Bedeutung zugesprochen (SCHARFENBERG, 2005). Es ist das Ziel dieser Lernorte Interesse und Motivation an naturwissenschaftlichen Fragestellungen zu wecken und naturwissenschaftliche Schlüsselkompetenzen auszubilden (FORSTER, LANGER, VETRVSKY & ANTON, 2011). Besonders der Lernort Schülerlabor schafft Zusatzangebote zum Unterricht und vor allem eine wichtige kompensatorische Funktion von experimentellen Anteilen (SCHARFENBERG, 2005). Beobachtungen im eigenen Schülerlabor haben gezeigt, dass die Teilnehmer mit Freude experimentieren. Weitere Arbeitsweisen, die zum Experimentieren gehören, wie zum Beispiel das Lesen der Versuchsanleitung, Aufräumen, Protokollieren und Deuten der Phänomene, wurden aber häufig abgelehnt. Hinweise der betreuenden Personen in Bezug auf diese Tätigkeiten wurden von den Schülern kommentiert und versucht zu umgehen. Da im Schülerlabor kein Leistungsdruck für die Teilnehmer besteht, schien diesen vor allem das praktische Arbeiten und das Beobachten eines Phänomens bedeutend. Die Experimente wurden mit wenig Blick auf Protokollvorlagen durchgeführt. Das Protokollieren von Versuchsergebnissen ist laut Literatur allerdings ein entscheidender Faktor zur Vernetzung von Theorie und Praxis. Abbildung 6 zeigt das Protokoll eines Teilnehmers aus dem Schülerlabor Nawi-Werkstatt.

UNIVERSITÄT KOBLENZ · LANDAU

NaWi-Werkstatt, 6. Juni 2012

„Kleine Kristalle mit großer Bedeutung“

Beobachtung:

Salz	Temperatur nach...			
	0 Sekunden (ohne Salz)	30 Sekunden	60 Sekunden	90 Sekunden
Natriumchlorid				
Kaliumnitrat		21,5	50,8	
Ammoniumnitrat				

Abbildung 6: Schülerprotokoll aus dem Schülerlabor Nawi-Werkstatt der Universität Koblenz-Landau.

Diese Beobachtung hat uns veranlasst, den Status Quo in Bezug auf das Protokollieren in anderen Schülerlaboren zu erheben (Teilstudie 1). Über den Bundesverband der Schülerlabore LELA (Lernort Labor) wurden im August 2012 die 327 Mitgliedslabore zum Protokollieren befragt. Über den Emailverteiler von LELA wurde ein mit GRAFSTAT erstellter Online-Fragebogen an die Leiter der Schülerlabore versendet. Der Fragebogen umfasste vier Fragen, welche in Abbildung 7 dargestellt sind. Je nach Angabe der Zielgruppe des Schülerlabors (Frage 1), erschien anschließend die Frage zur Nutzung der Protokolliermöglichkeiten dieser Altersgruppe (Frage 4). Besuchten mehrere Altersklassen das Schülerlabor, wurde diese Frage für jede Klassenstufen einzeln beantwortet. Zusätzlich wurden die Leiter der Schülerlabore über die bereitgestellten Materialien zum Protokollieren (Frage 2) und die Einstellung des Schülerlabors zu diesem Thema (Frage 3) befragt.

- *Welche Altersklassen besuchen Ihren außerschulischen Lernort?*
- *Welche Materialien benutzen die Teilnehmer zum Protokollieren?*
- *Fordern Sie die Teilnehmer explizit zum Protokollieren auf?*
- *Wie werden die Protokollmöglichkeiten von Schülern der Klassenstufe X Ihrer Einschätzung nach genutzt?*

Abbildung 7: Fragenkatalog der Fragebogenstudie zum Protokollieren in Schülerlaboren.

Die Einstufung der Altersklassen, die Materialien und die Aufforderung zum Protokollieren wurden über das Multiple-Choice-Antwortformat abgefragt. Der Fragebogen mit den Antwortformaten befindet sich im Anhang (Anhang I). Die Einschätzung zur Nutzung der Protokolliermöglichkeiten erfolgte über eine vierstufige Likert-Skala von „wenig genutzt“ bis „häufig genutzt“. Die Datenauswertung erfolgte mit GRAFSTAT. Da eine Mehrfachnennung der Altersklassen möglich war und sich die Anzahl der Altersklassen innerhalb der Schülerlabore unterschied, werden die Ergebnisse in relativen Häufigkeiten angegeben. Die Zahlen werden ohne Dezimalstelle angegeben, deshalb ergeben die Werte in der Summe nicht immer 100%.

Insgesamt nahmen n=87 Leiter der Schülerlabore an der Umfrage teil. Ein Schülerlabor wird meistens von mehreren Altersklassen besucht (Grundschule n=40, Orientierungsstufe n=42, 7.-10. Klasse n=67, Oberstufe n=66, Studenten n=12). 69% der Schülerlaborleiter gaben an, dass ihnen das Protokollieren wichtig ist. 11% ist der Prozess persönlich nicht wichtig, weshalb in diesen Schülerlaboren auch nicht zum Protokollieren aufgefordert wird. In 91% der

Schülerlabore wird das Protokollieren trotzdem ermöglicht, wohingegen in 9% der Schülerlabore generell kein Material heraus gegeben wird.

Abbildung 8 zeigt, dass Häufigkeit des Protokollierens in den Ausprägungen „selten“, „wenig“, „häufig“ und „sehr häufig“ für die jeweiligen Altersstufen, die die Schülerlabore besuchen. Die Angaben zur Grundschule zeigen, dass 32% der Schüler bei einem Schülerlaborbesuch wenig protokollieren. Bei einer Studentengruppe hingegen protokollieren nur 9% selten. Im Vergleich dazu wird bei den Studenten 36% häufig protokolliert, in der Sekundarstufe I protokollieren 9% viel bei einem Schülerlaborbesuch. 78% der Oberstufenschüler zeigen bei einem Schülerlaborbesuch eine mehr als angemessene Häufigkeit des Protokollierens, im Vergleich dazu zeigen 40% der Orientierungsstufenschüler die gleiche Häufigkeit auf.

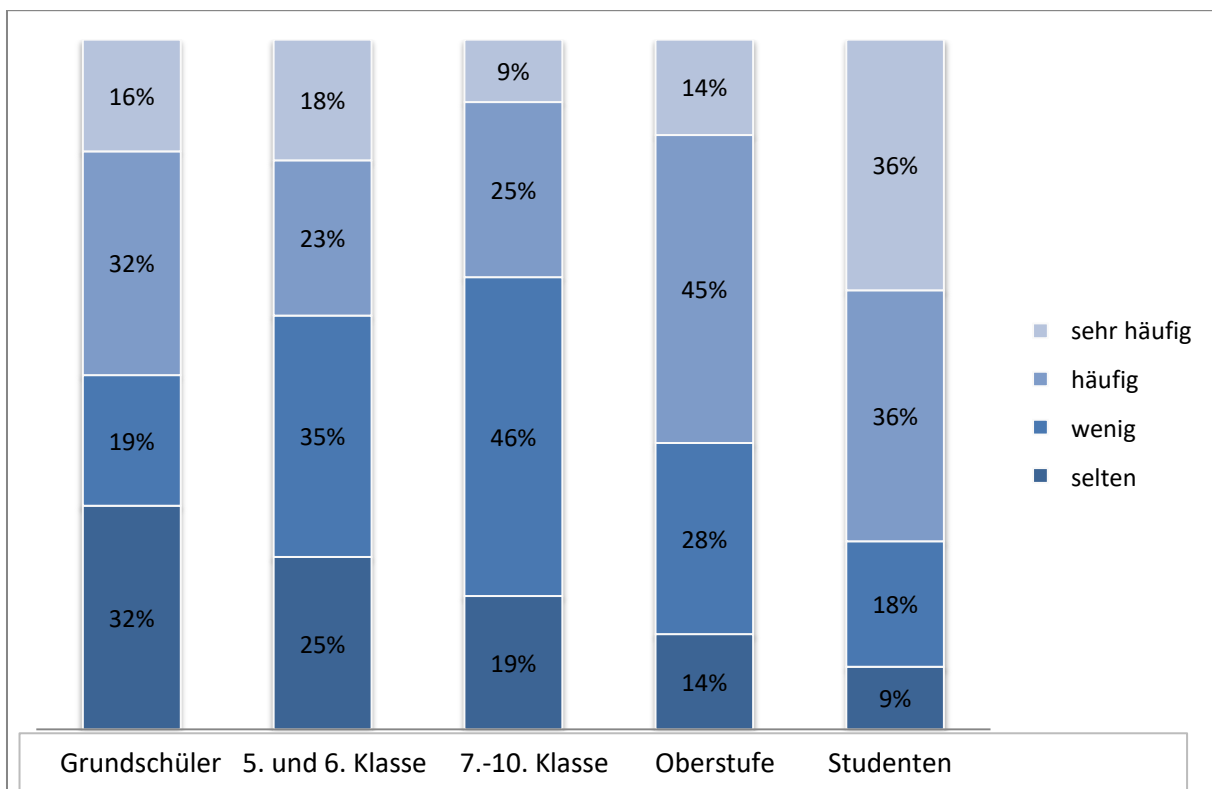


Abbildung 8: Relative Häufigkeiten der Antworten auf das Item „Wie werden die Protokollmöglichkeiten von Schülern der Klassenstufe X Ihrer Einschätzung nach genutzt?“ unterteilt nach Altersstufe.

Die Fragebogenstudie überprüft, ob allgemein in Schülerlaboren protokolliert wird und ob unsere Erfahrungen hinsichtlich der Motivation beim Protokollieren von anderen Schülerlaborleitungen geteilt werden. Dazu wurde zunächst untersucht, ob Schülerlaborleitungen die Schüler explizit zum Protokollieren auffordern. Dieses Item war notwendig, da die Arbeitsweisen vor Ort oft vom Ermessen der beteiligten Personen abhängen. Fast alle Probanden gaben an, dass sie das Protokollieren als wichtig erachten und es deshalb in den Ablauf des

Schülerlaborbesuches eingebunden ist. Dementsprechend werden den Schülern auch Materialien zum Protokollieren bereitgestellt. In den häufigsten Fällen handelt es sich dabei um ein leeres Protokollblatt. Einige Schüler erhalten für die Experimentierphasen auch vorstrukturierte Arbeitsmaterialien wie zum Beispiel Lückentexte oder Protokollvorlagen. Erachten die Leiter der Schülerlabore Protokollieren als unwichtig, werden keine Materialien verteilt. Für den Ablauf in diesen Schülerlaboren scheint ein Protokoll bedeutungslos und wird von den Teilnehmern nicht erwartet. Wahrscheinlich ist die Anstrengungsbereitschaft der Laborbesucher mit eigenen Materialien zu protokollieren unwesentlich. Dieser Sachverhalt muss bei der Interpretation der Häufigkeit des Protokollierens beachtet werden.

Die Abhängigkeit der Häufigkeit des Protokollierens und der Intensität des Protokollierens kann durch einen weiteren Aspekt veranschaulicht werden. Schüler aus der Grundschule und der Orientierungsstufe (5.-6. Klasse) protokollieren laut der Schülerlaborleitungen zu einem großen Teil nur angemessen. In den Klassenstufen 7-10 wird noch seltener das Protokollierangebot genutzt. Dementsprechend gering ist in diesem Alter die Motivation ein Protokoll anzufertigen. Bei den Schülerlaborteilnehmern der Oberstufe und Universität verdoppelt sich die Häufigkeit der Nutzung von Protokolliermöglichkeiten. Dieses Phänomen veranschaulicht, dass die Häufigkeit des Protokollierens von dem Alter der Teilnehmer abhängt. Die Häufigkeit des Protokollierens von Schülerlaborbesuchern ist über die Altersstufen hinweg nicht zu verallgemeinern. Es kann jedoch gezeigt werden, dass in Klassen mit jüngeren Schülern seltener protokolliert wird und deshalb eine Verbesserung des Protokollierens angestrebt werden sollte.

3 Das Protokoll als Element des naturwissenschaftlichen Erkenntnisprozesses

3.1 Theoretische Überlegungen – Bildungsstandards und Schreiben

Protokolle als Element der Bildungsstandards

Die von den Ergebnissen der großen Bildungsstudien TIMMS und PISA ausgelöste Bildungsdebatte führte im deutschen Bildungssystem zu einem Paradigmenwechsel von der Input- zur Output-Orientierung (FISCHER, KLEMM, LEUTNER, SUMFLETH, TIEMANN & WIRTH, 2003). Die Schüler sollen nach Ende ihrer Schulzeit dazu befähigt sein, naturwissenschaftliche Probleme zu lösen und an gesellschaftlichen Diskussionen teilnehmen zu können (KMK, 2005a, 2005b, 2005c). Dieses Ziel soll durch die Vermittlung der vier grundlegenden Kompetenzen Fachwissen, Erkenntnisgewinnung, Kommunikation sowie Bewertung erreicht werden. Der Kompetenzbereich Fachwissen umfasst das Wissen über chemische Phänomene und das Verständnis grundlegender chemischer Begriffe, Prinzipien und Gesetzmäßigkeiten (KMK, 2005b). Experimentelle Untersuchungsmethoden und fachbezogene Denk- und Arbeitsweisen sind dem Kompetenzbereich Erkenntnisgewinnung zuzuordnen (ebd.). Im Bereich Kommunikation findet ein fachbezogener Informationsaustausch durch die Verknüpfung von Alltags- und Fachsprache statt (ebd.). Kompetenzen im Bereich der Kommunikation sind notwendig, um Phänomene zu erklären und in geeigneter Form darzustellen (ebd.). Im Kompetenzbereich Bewertung werden chemische Sachverhalte und Kontexte reflektiert (ebd.). Die Fähigkeit ein Protokoll anzufertigen zu können, wird sowohl im Kompetenzbereich Erkenntnisgewinnung (Abkürzung E) als auch im Kompetenzbereich Kommunikation (Abkürzung K) gefordert:

Die Schüler

- führen qualitative und einfache quantitative experimentelle und andere Untersuchungen durch und protokollieren diese (E 3),
- erheben bei Untersuchungen, insbesondere in chemischen Experimenten, relevante Daten oder recherchieren sie (E 5),
- finden in erhobenen oder recherchierten Daten, Trends, Strukturen und Beziehungen, erklären diese und ziehen geeignete Schlussfolgerungen (E 6),

- protokollieren den Verlauf und die Ergebnisse von Untersuchungen und Diskussion in angemessener Form (K 6),
- dokumentieren und präsentieren den Verlauf und die Ergebnisse ihrer Arbeit situationsgerecht und adressatenbezogen (K 7) (KMK, 2005b).

Der Kompetenzbereich Erkenntnisgewinnung ist in seinen Teilkompetenzen aktuell am detailliertesten operationalisiert. Nach WELLNITZ ET AL. (2012) wird der Kompetenzbereich Erkenntnisgewinnung in die Teilbereiche „Naturwissenschaftliche Untersuchungen“, „Naturwissenschaftliche Modellbildung“ und „Wissenschaftstheoretische Reflexion“ unterteilt. Im Teilbereich „Naturwissenschaftliche Untersuchungen“ finden sich vor allem Begriffe und Methoden von „Scientific Inquiry“ wieder. Unter diesem Begriff werden die Durchführung und das Verständnis zur Lösung naturwissenschaftlicher Probleme im internationalen Sprachraum gefasst (WELLNITZ ET AL., 2012). Die Schüler sollen beispielsweise Arbeits- und Auswertungsmethoden beherrschen und anwenden können (KMK, 2005a, b, c). Es ist demnach die Aufgabe des naturwissenschaftlichen Unterrichts grundlegende naturwissenschaftliche Arbeitsweisen zu erlernen (FISCHER ET AL., 2003). Die Methoden der naturwissenschaftlichen Untersuchungen zeichnen sich durch den Erkenntnisprozess aus. In diesen Teilstrukturen des Erkenntnisweges wird auch die Fähigkeit ein Protokoll anzufertigen gefordert. Die Abbildung 9 stellt diese Operationalisierung ergänzend dar:

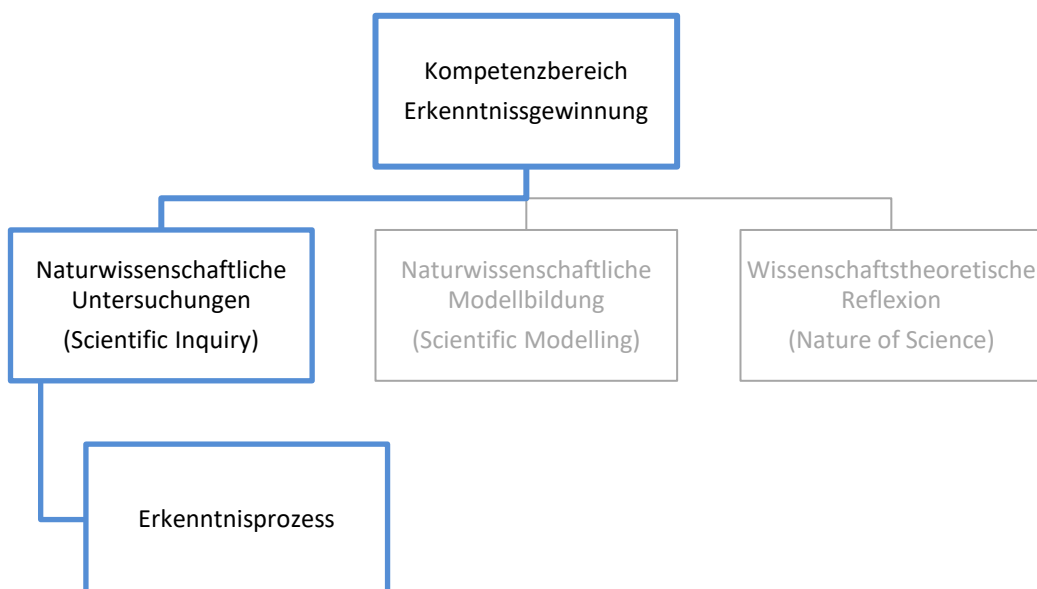


Abbildung 9: Operationalisierung des Kompetenzbereichs Erkenntnisgewinnung und Einordnung des Erkenntnisprozesses.

Ein zentraler Aspekt dieser Denk- und Arbeitsweise ist das Experiment (PRENZEL & PARCHMANN, 2003). Das Experimentieren stellt einen Prozess dar, welcher in eigenständige Teilprozesse unterteilt werden kann (WELLNITZ ET AL., 2012). Das Experimentieren ist dadurch nicht als eine einzelne Fähigkeit zu verstehen, sondern verlangt von dem Lernenden mehrere Teilfähigkeiten (ebd.). Dieser Erkenntnisprozess wird in der Literatur durch unterschiedliche Arbeitsschritte dargestellt und diskutiert. Die unterschiedlichen Strukturierungen eines Experimentes einzelner Autoren zeigt eine Zusammenstellung von EMDEN (2011) (Abbildung 10). Die Arbeitsweisen sind verschieden benannt, ähneln sich jedoch in der Bedeutung. Deshalb lassen sich die Schritte in drei Phasen zusammenfassen (ebd.). Der erste Teilprozess stellt zu meist das Bilden einer Hypothese oder die Auseinandersetzung mit dem Problem oder der Fragestellung (I). In anderen Auffassungen hingegen bilden beispielsweise die Planung und das Design den ersten Teilschritt. Darauf folgt in der Regel eine Überprüfung durch die Durchführung eines Experiments (II) dar. Dieser Teilprozess ist in allen Experimentierstrukturen wiederzufinden. Jedoch variieren die Benennungen von „Experimentieren“, „Durchführung“ bis hin zu „Prüfen von Hypothesen“. Abschließend wird meist der Teilprozess der Auswertung oder Schlussfolgerung genannt (III). Die Experimentierstrukturen variieren von mindestens drei bis zu maximal fünf Teilprozesse. Die von EMDEN (2011) postulierte Dreiteilung beschreiben KLAHR & DUNBAR (1988) mit dem „Scientific Discovery as Dual Search“ Modell (SDDS-Modell) den naturwissenschaftlichen Erkenntnisweg mit den drei Komponenten „Suche im Hypothesenraum“, „Testen von Hypothesen“ und „Auswertung experimenteller Evidenz“. Das Experimentieren wird nach dieser Theorie als Problemlösestrategie für ein Problem angesehen, der durch die Suche im Hypothesensuchraum (Space of Hypothesis) und dem Experimentiersuchraum (Space of Experiments) zu einer Schlussfolgerung gelangt. Der Hypothesensuchraum setzt sich aus den aufgestellten Hypothesen und der Analyse der Ergebnisse zusammen. Der Experimentiersuchraum besteht aus den daraus resultierenden möglichen Experimenten. Eine Hypothese wird gebildet und anschließend ein Experiment gesucht, welches Ergebnisse liefert, um in Bezug auf die Hypothese interpretiert zu werden (KLAHR, 2000). In keinem dieser Schritte findet sich explizit der Schritt des Protokollierens wieder, der von den Schülern in diesem Kompetenzbereich jedoch gefordert wird. MAYER (2007) benennt vier Teilprozesse der naturwissenschaftlichen Erkenntnisgewinnung, die als eigenständig identifiziert wurden: (1) Fragestellung formulieren, (2) Hypothesen generieren, (3) Untersuchung planen und (4) Daten analysieren und schlussfolgern. Er trennt den Punkt

der Planung in zwei Schritte und stellt somit diese Teilschritte in den Fokus des Erkenntnisprozesses. Um einen Teilschritt von den anderen Schritten trennscharf darstellen zu können, ist eine genaue Operationalisierung des Erkenntnisprozess notwendig.

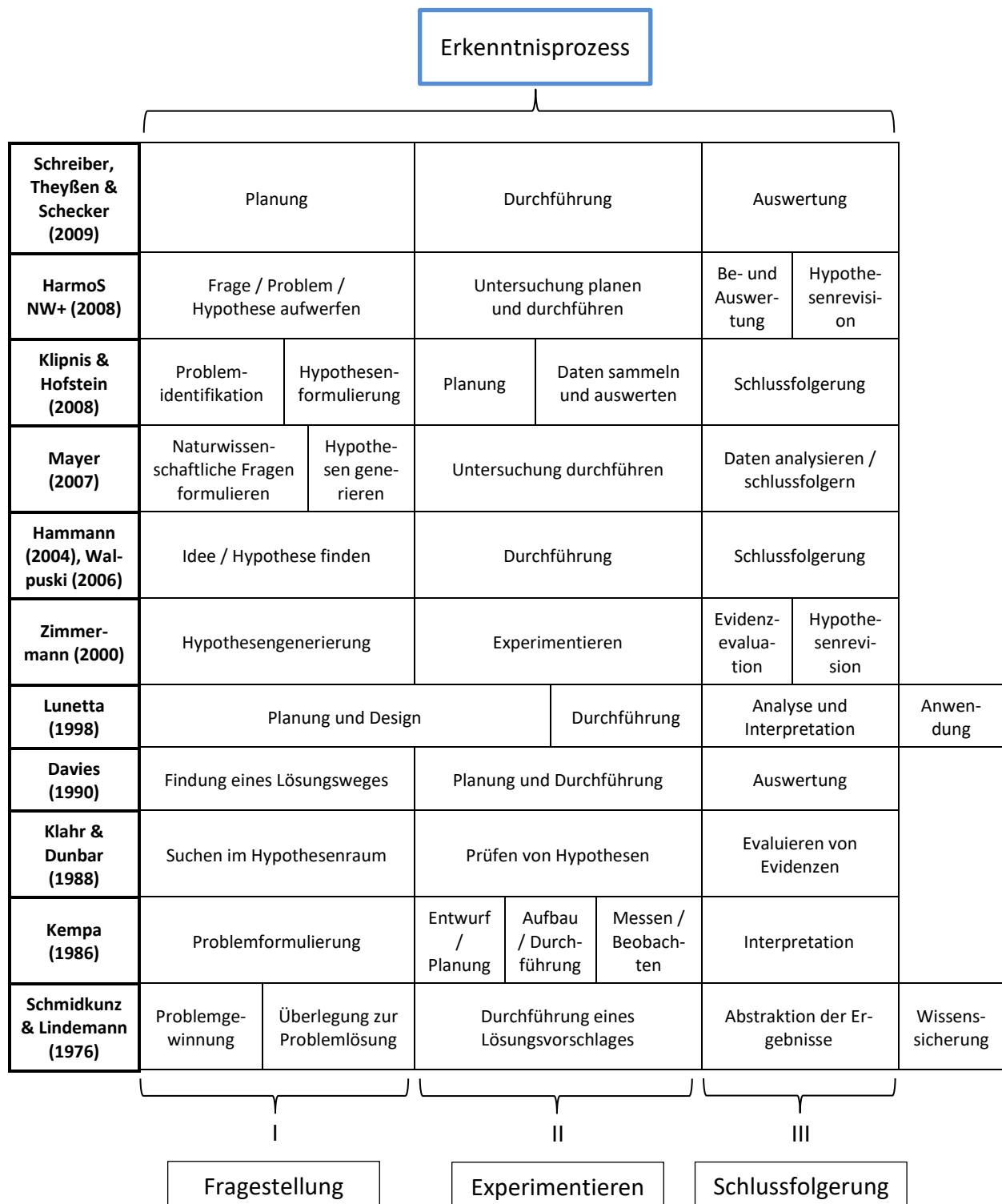


Abbildung 10: Strukturierung der Teilprozesse des Erkenntnisweges nach unterschiedlichen Autoren. Quelle: modifiziert nach EMDEN (2011).

Die beschriebene Teilprozesse wie beispielsweise „Auswertung“, „Daten sammeln und auswerten“, „Daten analysieren / Schlussfolgern“ oder „Abstraktion der Ergebnisse“ beinhalten die Sicherung und Auswertung der erhobenen Daten. Diese Sicherung erfolgt in der Regel über die Erstellung eines Protokolls. Damit ist nicht zwangsläufig das klassische Versuchsprotokoll gemeint, indem beispielsweise Versuchsaufbau und Versuchsdurchführung für eine Reproduzierbarkeit eindeutig beschrieben werden müssen (KRAUS & STEHLIK, 2008). In diesem Sinne des Protokolls geht es lediglich um das Festhalten der Versuchsergebnisse, sowie deren Auswertung. Durch die Definition des Protokolls ist der Ausdruck „Daten aufschreiben und auswerten“ am zutreffendsten. Um diesen Teilschritt in dem Erkenntnisprozess hervorzuheben, bietet sich eine Vierteilung an der von EMDEN (2011) zusammengefassten Dreiteilung an (Abbildung 11).

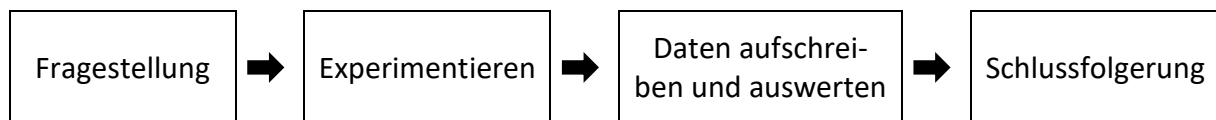


Abbildung 11: Angepasste Struktur der Teilschritte des Erkenntnisprozesses.

Der Umgang mit Daten und deren Evaluation ist elementarer Bestandteil wissenschaftlichen Denkens (KLAHR, 2000). In der Schule werden zum Erlernen des Umgangs mit Daten in der Regel Daten aus Experimenten verwendet (PFIELER, PRIEMER & UPMEIER ZU BELZEN, 2015). Laut HUG & MCNEILL (2008) werden Daten, die während eigener Untersuchungen gewonnen werden, als Primärdaten bezeichnet. Stammen die Daten aus der Arbeit anderer, werden sie Sekundärdaten genannt. In welcher Form die Daten erhoben und festgehalten werden, wird nicht explizit erwähnt. Eine Weitergabe der Daten durch ein Protokoll scheint jedoch sinnig. Protokolle werden vorrangig mit Primärdaten aus eigenen Untersuchungen angefertigt.

Sprache ist mit den Kompetenzen der naturwissenschaftlichen Untersuchungen eng verknüpft (BEERENWINKEL, HEFTI, LINDAUER & SCHMELLENTIN, 2016). Die Verknüpfung der Kompetenzbereiche Erkenntnisgewinnung und Kommunikation wird durch das Experimentieren und dem damit verbundenen Erzeugen von Daten erreicht. Um den Kenntnisstand der Schüler im Kompetenzbereich Kommunikation erfassen zu können, ist eine Operationalisierung dieser Kompetenz notwendig. Durch eine Verknüpfung von Modellen der Kommunikationswissenschaften und bisheriger Arbeiten der naturwissenschaftsdidaktischen Forschung leitet KOBOW (2015) drei Kompetenzteilbereiche und drei Aspekte der Kommunikationskompetenz ab (Tabelle 4). Die Teilbereiche „Informationen erschließen“, „Informationen weitergeben“

und „Argumentieren“ werden innerhalb des Kommunikationsprozess gemeinsam genutzt und lassen sich deshalb auch nicht trennscharf voneinander unterscheiden (ebd.). Auch die drei Aspekte der Kommunikationskompetenz „Sprache / Fachsprache“, „Darstellungsform“ und „Adressatenbezug / Sachbezug“ können gleichzeitig genutzt werden und sind deshalb miteinander verknüpft (ebd.). In einem Protokoll können alle Teilbereiche und Aspekte der Kommunikation umgesetzt werden: Informationen in Form von Daten werden über die Sprache und unterschiedliche Darstellungsformen erschlossen. Das Protokoll kann zum Austausch an Lernende und Lehrende weitergeben werden und somit wird ein möglicher Adressatenbezug hergestellt.

Tabelle 4: Operationalisierung des Kompetenzbereich Kommunikation in Kompetenzteilbereiche und Aspekte. Quelle: Kobow (2015).

Kommunikation			
Kompetenzteilbereiche	Informationen erschließen	Information weitergeben	Argumentieren
Aspekte	Sprache / Fachsprache		
	Darstellungsformen		
	Adressatenbezug / Sachbezug		

Das Protokollieren stellt einen wichtigen Teilbereich der Kompetenzbereiche Erkenntnisgewinnung und Kommunikation dar. Der Kompetenzbereich Kommunikation der Bildungsstandards zielt auf eine Förderung der Sprache und der Kommunikationsprozesse im naturwissenschaftlichen Unterricht ab (KMK, 2005b). Lehren und Lernen ist ohne die Nutzung von Sprache nicht möglich (SUMFLETH, KOBOW, TAUNALI & WALPUSKI, 2013). Deshalb sind sprachliche Kompetenzen für ein erfolgreiches Lernen in jedem Fach notwendig (BECKER-MROTZEK, SCHRAMM, THÜRMANN & VOLLMER, 2013). Sprachliche Bildung erfolgt zwar alltagsintegriert, ist aber eine Aufgabe der Bildungsinstitutionen und stellt ein gezieltes Element in jeder Unterrichtsvorbereitung dar (RALLE, 2015). Die Durchführung von Experimenten verbindet dadurch auch die sprachliche Entwicklung mit dem Ziel der naturwissenschaftlichen Erkenntnisgewinnung (BUSCH & RALLE, 2013). LÜCK (2008) bezeichnet das Experimentieren und das damit verbundene Bedürfnis sich sprachlich auszudrücken als Anlass für Sprache und Sprachförderung im naturwissenschaftlichen Unterricht. Das Erstellen von Protokollen ist während des Experimentierens die schwierigste Phase der Fachsprache (BUSCH & RALLE, 2013).

Zur Überprüfung von Kompetenzen gilt es die einzelnen Teilbereiche des Erkenntnisprozess zu hinterfragen und genauer zu definieren, um diese bei den Schülern bestmöglich entwickeln zu können. Dabei stellt sich die Frage, ob alle vorgeschriebenen Kompetenzen auch zu einem Wissenserwerb beitragen. Deshalb ist es notwendig einzelne Teilschritte und Methoden der Erkenntnisgewinnung genauer zu untersuchen. Will man diesen Prozess bei Schülern erforschen und mit Hilfe von Lernumgebungen unterstützen, dann müssen dazu einzelne, bei der Erkenntnisgewinnung benötigte Kompetenzen erfasst werden. Von besonderer Bedeutung erscheint in diesem Zusammenhang die Fähigkeit zu sein, Vorgehensweisen und Ergebnisse mit Hilfe von passgenauen Repräsentationen so zu protokollieren, dass die entstehenden Dokumente sich gewinnbringend für weitere Problemlöse-, Erkenntnis- und Lernprozesse nutzen lassen. In der fachdidaktischen Forschung gibt es bisher keine Studien, in denen explizit der Teilschritt der Datenaufnahme und dessen Bedeutung untersucht wurde. Um Kompetenzen Lernender erfassen und gezielt fördern zu können, ist die Entwicklung theoriebasierter Kompetenzmodelle von großer Bedeutung (HELMKE & HOSENFELD, 2004). Kompetenzstrukturmodelle können durch die Differenzierung von Teilfähigkeiten einer Kompetenz die Anforderungen beschreiben (KLIEME, AVENARIUS, BLUM, DÖBRICH, GRUBER & PRENZEL, 2003). Um die Anforderungen an die Bildungsstandards empirisch zu prüfen, müssen Kompetenzmodelle formuliert werden, die die Fähigkeiten von Schülern beschreiben (SCHECKER & PARCHMANN, 2006).

Bedeutung des Schreibens für die Naturwissenschaften

Ein möglicher Grund für die mangelnde Beliebtheit der Protokolle liefert die Tätigkeit des Schreibens. Die Verschriftlichung und Aufbereitung von Ergebnissen ist ein wesentlicher Teil beim Protokollieren, mit dem Ziel eine Schlussfolgerung zu ziehen (KMK, 2005b; SCHREIBER, THEYSEN & SCHECKER, 2009; WITTECK & EIKS, 2004). Viele Schüler haben erhebliche Schwierigkeiten beim Schreiben (MOLL, 2003; PRECHTL, 2008). Alleine die Tätigkeit und das Finden der richtigen Wörter fällt nicht nur Schülern der Orientierungsstufe schwer (MOLL, 2003; STEINHOFF, 2003). Ungeübte Schreiber sind schnell damit überfordert ihr Wissen in Worte zu fassen (HÜBNER, NÜCKLES & RENKL, 2007). Auch wegen der fehlenden Selbstbestimmung (Übernehmen von Merksätzen etc.) in dieser Form ist das Schreiben die unbeliebteste Tätigkeit im naturwissenschaftlichen Unterricht (BERGLER, 2009). Allerdings unterstützt die schriftliche Auseinandersetzung mit dem Lernstoff die Anwendung wichtiger kognitiver und metakogni-

tiver Lernstrategien (HÜBNER ET AL., 2007; BERGELER & POSPIECH, 2007). Es gilt daher die Einstellung gegenüber dem Protokollieren zu verbessern. Dies kann durch die Unterstützung der Schüler in ihrer Tätigkeit erfolgen.

Das Schreiben im Fach nimmt die Rolle der schriftsprachlichen Kompetenzen für das fachliche Lernen ein und wird von verschiedenen Autoren betont (zum Beispiel KEYS, 1994; 1999; BERGELER, 2009; STEPHANY, LINNEMANN & BECKER-MROTZEK, 2013; SCHUMACHER & ROTH, 2014). Schreiben ist ein komplexer Prozess, bei dem Wissen nicht nur dargestellt, sondern auch neues Wissen aufgebaut und geordnet wird (vgl. STORK, 1993). Auf der internationalen Ebene hat sich die Forschungseinrichtung „writing to learn science“ bereits seit den 1970er Jahren etabliert (HOLLIDAY, YORE & ALVERMANN, 1994; RIVARD, 1994). Beginn dieser Bewegung war die Arbeit von Britton (NIESWANDT, 1998). Für BRITTON (1970) haben Schreibtätigkeiten der Schüler den Zweck, dass der Lehrer das Wissen der Schüler überprüfen kann und zusätzlich die Schüler unterstützt, Wissen zu konstruieren und ihre Interpretation der Inhalte zu präsentieren. DURST & NEWELL (1989) kommen zu der Erkenntnis, dass das Schreiben von Notizen, Zusammenfassungen und Antworten von Verständnisfragen zu einem lesenden Text einen ähnlichen Effekt haben, und dass es hilft, Wissen zu überprüfen, zu festigen und zu behalten. Nach NIESWANDT (1998) wird in anderen Forschungen von FELLOWS (1994), GLYNN & MUTH (1994), KEYS (1994) und SANTA & HAVENS (1991) das Schreiben als ein Weg gesehen, der das Vorwissen mit neuem Wissen verknüpft, naturwissenschaftliche Konzepte entwickelt und erklärt, Denkprozesse offen gezeigt und dem Lernenden dabei hilft Informationen aus Experimenten und Büchern zu verarbeiten. In diesen Schreibprojekten werden das darstellende Schreiben (*expository writing*) für das Beschreiben und Deuten experimenteller Beobachtungen und das expressive Schreiben (*expressive writing*) für informelles Schreiben verwendet (NIESWANDT, 1998). Wenn das Schreiben nach diesen Studien so intensiv auf den Denkprozess des Lernenden einwirkt, ist es nicht verwunderlich, dass viele Studien aufzeigen, dass unter bestimmten Voraussetzungen eine positive Wirkung auf den Lernerfolg durch das Schreiben im naturwissenschaftlichen Unterricht erzielt werden kann (vgl. WALLACE, HAND & PRAIN, 2004; YORE, 2003; PRAIN & HAND, 1996). Jedoch geht es in diesen Studien primär um das Anfertigen von Fachtexten und nicht spezifisch um das Schreiben der Protokolle.

Das Schreiben im naturwissenschaftlichen Unterricht fand im deutschsprachigen Raum erst später die notwendige Beachtung (STORCK, 1993). Die Arbeiten der Schweizer GALLIN & RUF (1991, 1993, 1995) stellen in dieser Hinsicht Pionierarbeit im deutschsprachigen Raum dar.

Die Autoren zeigen auf, dass das Arbeiten im mathematischen Problemlöseprozess durch die schriftliche Darstellung zwar verlangsamt wird, dies jedoch mit einer regelmäßigen Leistungssteigerung verbunden ist. Hinter der Idee von *learning by writing* steckt die kognitive Verarbeitung des Lerninhaltes durch Anwendung von Lernaktivität, wie zum Beispiel Identifizierung von Ideen, Ausarbeitung und kritische Reflexion (BERTHOLD ET AL., 2007). Durch solche Lernaktivitäten werden das Verständnis und die Aufbewahrung der Inhalte erlernt (vgl. PINTRICH, SMITH, GARCIA, & MCKEACHIE, 1991). Darüber hinaus können durch metakognitive Aktivitäten Ungereimtheiten und Lücken im Text und somit auch im Wissen leichter vom Lernenden entdeckt werden (*reflecting by writing*) (BERTHOLD ET AL., 2007). Nach DÖRFLER (2003) können Schüler mit Hilfe ihrer Versprachlichung den Lernprozess verarbeiten und zusätzlich mit Lernenden und Lehrenden über den Lernprozess kommunizieren. Das Protokoll ist somit ein kommunikatives Ausdrucksmittel für die Schematisierung von Lernprozessen und ihren Ergebnissen (ebd.). Es wird deshalb davon ausgegangen, dass dadurch die Lerninhalte tiefer durchdrungen werden und ein differenzierter Blick auf die wichtigen Zusammenhänge erreicht werden kann (ebd.). BERGELER (2009) gibt Aufschluss darüber, dass sich das Schreiben als Lernmethode für den naturwissenschaftlichen Unterricht eignet. Wissenskonstruktion findet beim eigenständigen Schreiben statt, da der Schreiber die einzelnen Sachverhalte für die Repräsentation in einem sinnvollen Text verknüpfen und an die gegebene Situation anpassen muss, um das Textziel zu erfüllen (ebd.). Die Ergebnisse zeigen, dass der Lernerfolg der Schüler durch die Schreib-Lern-Methode steigt, sich die Methode jedoch nicht für eine Motivationsförderung eignet (BERGELER & POSPIECH, 2007). In der aktuellen Forschung zeigt sich eine Tradition der Integration von sprachlichen und mathematischen Ressourcen (PRAIN & TYTLER, 2013). Es wird zunehmend anerkannt, dass die Schüler lernen müssen, zu interpretieren und zu konstruieren sowie wissenschaftliche Konzepte, Verfahren, Ansprüche und Erkenntnisse darzustellen (ebd.).

3.2 Empirische Studie zum Schreiben von Protokollen–Darstellungsart und Einflussfaktoren

Die Forschungsinitiative II der Universität Koblenz-Landau im Bereich der Bildungswissenschaften zielte darauf ab, eine systematische Vernetzung des schulischen Lernorts Unterricht mit außerschulischen Lernorten unter der Perspektive der Fachdidaktik, Pädagogik und Psychologie herzustellen. Dabei ging es darum, spezifische Synergieeffekte für Lernprozesse zu schaffen und zu nutzen, die sich sowohl auf die Aneignung von fachlichem Wissen als auch auf den Transfer dieses Wissens auf neue Anwendungssituationen oder neue Domänen positiv auswirken. Als ein wesentliches Element zur Förderung und Entwicklung wissenschaftlichen Denkens und Arbeitens wurde die Fähigkeit identifiziert, Arbeitsprozesse geeignet zu reflektieren und deren Ergebnisse mit Hilfe von passgenauen Repräsentationen so festzuhalten, dass die entstehenden Dokumente sich gewinnbringend für weitere Problemlöse-, Erkenntnis- und Lernprozesse nutzen lassen. Diese Prozesse wurden vor allem unter den Schwerpunkten Protokollieren und Darstellen sowie Erfassen und Identifizieren betrachtet. Die Forschungsinitiative bestand aus mehreren Teilprojekten. Seitens der mathematisch-naturwissenschaftlichen Fachdidaktikern wurde der Schwerpunkt auf das Protokollieren und Darstellen gelegt. Durch diese Kooperation wurde das Schreiben von Protokollen durch unterschiedliche Fachdisziplinen beleuchtet und dadurch ein fachübergreifender Kontext geschaffen. Die Untersuchung fand jeweils im Schülerlabor der Mathematik „Mathe ist mehr!“ und Chemie „Nawi-Werkstatt“ statt. Für diese Teilstudie galt es sich der Theorie des Protokolls und dem Prozess des Protokollierens im naturwissenschaftlichen Bereich auf explorative Weise zu nähern. Viele der folgenden Kriterien von Design und Intervention ergaben sich durch die Kooperation im Forschungsverbund. Es galt die Annahme, dass sich durch diese Kriterien die Protokolle und die Tätigkeit des Protokollierens der Schüler voneinander unterscheiden. Im Rahmen dieser Arbeit wurden deshalb folgende Forschungsfragen formuliert:

➤ Forschungsfrage 1.1

Durch welche design- und interventionsbedingten Kriterien wird das Produkt Protokoll beeinflusst?

Obwohl alle Schüler am gleichen Inhalt arbeiten, können unterschiedliche Produkte. Die Unterscheidungsmerkmale der Protokolle werden deduktiv als auch induktiv bestimmt. Es stellt sich jedoch die Frage, durch welche Merkmale unterschiedliche Protokolle entstehen. Deshalb stellen die Intervention und das Design bewusste Einflüsse auf das Protokoll dar.

➤ Forschungsfrage 1.2

Durch welche design- und interventionsbedingten Kriterien wird die Tätigkeit des Protokollierens beeinflusst?

Es gilt nicht nur das Protokoll als Endprodukt isoliert zu betrachten, sondern auch den Prozess der Entstehung zu berücksichtigen. Diese beiden Teilbereiche bedingen sich gegenseitig. Deshalb ist auch die Beeinflussung der Kriterien des Designs und der Intervention auf die Tätigkeit des Protokollierens auf Unterscheidungsmerkmale zu untersuchen. Die derzeitige Literatur gibt wenig Aufschluss über den Prozess des Protokollierens, deshalb ist eine induktive Herangehensweise schlüssig.

Erarbeitungsprotokolle

Im Rahmen dieser Teilstudie wurde der Begriff „Erarbeitungsprotokoll“ eingeführt (vgl. ENGL ET AL., 2015). Die Kooperation der Fachdisziplinen erforderte eine neue Begrifflichkeit für den Protokollbegriff, da durch die Fachliteratur unterschiedliche Definitionen (vgl. Kapitel 2.1) vorhanden waren. In Erarbeitungsprotokollen werden Vorgehensweisen und Ergebnisse während eines selbständigkeitsorientierten Erarbeitungsprozesses festgehalten. Sie unterscheiden sich somit von den Begriffen wie „Versuchsprotokoll“ oder auch „Lernprotokoll“, da keine Vorgaben für das Protokoll gemacht werden. Trotzdem steht ein Experiment oder eine Aufgabe im Fokus und nicht der Lernprozess. Die Materialien und Durchführung sind in der Versuchsanleitung bereits gegeben und somit sind nur die Teilbereiche Beobachtung und Ergebnisse zu notieren. Im weiteren Verlauf dieser Arbeit ist das Protokoll mit der Definition des Erarbeitungsprotokolls gleichzusetzen.

Design und Stichprobe

Zur Beantwortung der Forschungsfragen wurde ein Projekt generiert, welches von April 2012 bis März 2013 stattfand (vgl. Abbildung 12). Zunächst wurden die eingesetzten Unterrichts- und Testmaterialien konzipiert, pilotiert und anschließend überarbeitet. Als Untersuchungsdesign wurde die Methode der Vergleichsstudie mit zwei mal zwei Interventionsgruppen und drei Messzeitpunkten gewählt. Jede Gruppe besuchte drei Wochen für je einen Termin pro Woche das Schülerlabor Nawi-Werkstatt und ebenso drei Wochen für je einen Termin pro Woche das Mathematik-Labor „Mathe ist mehr!“. Der Pre-Test fand in der Woche vor der Intervention statt. Der Post-Test wurde jeweils in der Woche nach der Intervention im jeweiligen Schülerlabor durchgeführt. Da nicht alle Klassen gleichzeitig die sechswöchige Intervention durchführten, erstreckte sich der zeitliche Rahmen der Durchführung über zwei Monate. Drei Monate nach dem Post-Test fand der Follow-Up-Test statt. Eine Woche nach dem Follow-Up-Test wurden zehn zufällig ausgewählte Schüler der Vergleichsstudie interviewt.

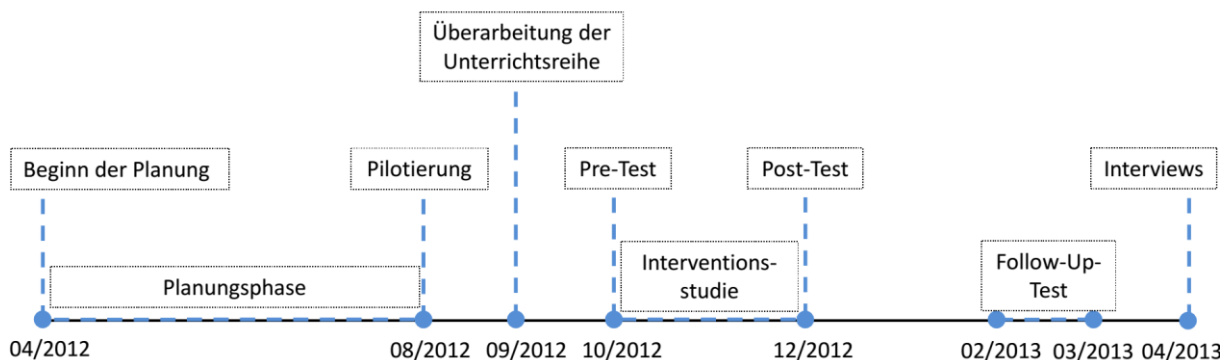


Abbildung 12: Zeitlicher Verlauf der empirischen Studie.

Die 196 Probanden (sechs Klassenverbände) dieser Teilstudie waren Schüler der sechsten Klassenstufe von zwei Gymnasien aus Landau (Max-Slevogt-Gymnasium und Otto-Hahn-Gymnasium) mit einem gemittelten Alter von 11,1 Jahren. Wir entschieden uns bewusst für die Durchführung mit einer sechsten Klasse und eine thematische Einführung in das Themenfeld 7 des Fachs Naturwissenschaften, da die Schüler zu diesem Zeitpunkt nach Lehrplan noch keine Einführung in das Protokollieren erhalten hatten. Eine Einführung des Protokollierens durch den Lehrer würde zu einer unterschiedlichen Vorerfahrung der Schüler führen. Nach Absprache mit den Lehrern fand zu diesem Zeitpunkt noch keine Unterrichtseinheit zum Protokollieren statt. Die Schülergemeinschaft setzte sich aus 42% Mädchen und 58% Jungen zusammen. Davon wiesen 17% einen Migrationshintergrund auf. Ihre Leistung im Fach Naturwissenschaften lag im Schnitt bei der Schulnote 2,3. Allerdings willigten 2% der

Eltern bei der Einverständniserklärung für die Freigabe der demografischen Daten nicht ein, wodurch fehlende Werte zustande kamen. Die Studie entspricht dem quasiexperimentellen Vergleichsgruppendesign. Um die Lernvariable (Merkmale und Voraussetzungen von Schülern) möglichst zu kontrollieren, fand die Einteilung der Kontroll- und Experimentalgruppe durch eine Teilung des Klassenverbandes statt (vgl. Abbildung 13). Die Gruppeneinteilung wurde aus organisatorischen Gründen von den Lehrern vorgenommen und ist deshalb quasi-randomisiert. Die beiden Interventionsgruppen arbeiteten jeweils in getrennten Räumen. Die Einteilung wurde über die drei Besuche hinweg beibehalten.

Doppelstunde 1	Doppelstunde 2	Doppelstunde 3
Klasse 1: Experimentalgruppe		
Klasse 1: Kontrollgruppe		
Klasse 2: Experimentalgruppe		
Klasse 2: Kontrollgruppe		
Klasse 3: Experimentalgruppe		
Klasse 3: Kontrollgruppe		
Klasse 4: Experimentalgruppe		
Klasse 4: Kontrollgruppe		
Klasse 5: Experimentalgruppe		
Klasse 5: Kontrollgruppe		
Klasse 6: Experimentalgruppe		
Klasse 6: Kontrollgruppe		

Abbildung 13: Übersicht der Einteilung in die Untersuchungsgruppen; Experimentalgruppe (EG), Kontrollgruppe (KG).

Eine zweite Kreuzung der Gruppen wurde durch die Reihenfolge der Besuche des jeweiligen Schülerlabors vorgenommen. Drei Klassen besuchten zuerst die Einheit der Nawi-Werkstatt (BeginnN), während die anderen drei Klassen anfangs im Mathematik-Labor tätig waren (BeginnM). Nach drei Doppelstunden wechselten die Gruppen die außerschulischen Lernorte. Dieses zweite Vergleichsgruppendesign ist in Abbildung 14 erläutert. Die Einteilung in Kontroll- und Experimentalgruppe blieb über diesen Vergleich hinweg bestehen.

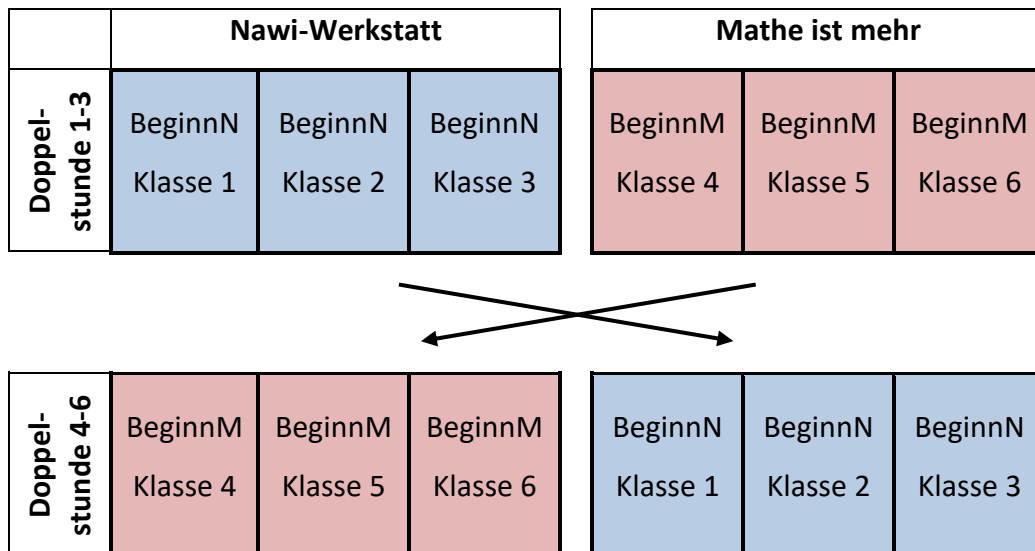


Abbildung 14: Übersicht des Vergleichsgruppendesigns, BeginnN = Beginn Nawi-Werkstatt, BeginnM = Beginn „Mathe ist mehr“.

Innerhalb der Klasse arbeiteten die Schüler in Kleingruppen von drei bis vier Personen. Jede Klasse wurde je nach Klassenstärke in sieben bis acht Kleingruppen eingeteilt, drei oder vier Gruppen in der Kontrollgruppe und vier Gruppen in der Experimentalgruppe. Auch diese Gruppeneinteilung blieb über beide Schülerlabore hinweg bestehen.

Interventionsmaßnahmen

Die Strukturierung der Interventionsmaßnahme war in der Nawi-Werkstatt durch die Bearbeitung von sieben Experimenten gegeben. In der ersten Doppelstunde wurden zwei Experimente, in der zweiten Doppelstunde drei Experimente und in der dritten Doppelstunde zwei Experimente bearbeitet. Die Schüler hatten jeweils 90 Minuten Bearbeitungszeit für zwei beziehungsweise drei Experimente. Eine Fragestellung konnte mit Hilfe der vorhandenen Materialien und der Versuchsbeschreibung bearbeitet werden (siehe Kapitel 3.2.2). Der zeitliche Rahmen, die thematische Fragestellung, Material und Anleitung wurden in allen Gruppen gleich gestaltet. Die sich unterscheidenden Interventionsmaßnahmen wurden während der Bearbeitungszeit der Experimente eingesetzt.

In der Kontrollgruppe erhielten die Schüler zu Beginn der Studie die Aufforderung ihre Ergebnisse der Experimente zu protokollieren und in ihrem Forscherheft darzustellen. Dazu stand ihnen ein Klemmbrett mit einem Stift und dem jeweiligen Arbeitsheft für das Experiment zur Verfügung. Im Arbeitsheft war dafür ein Bereich vorgesehen und mit einem Symbol entsprechend gekennzeichnet (Abbildung 15).



Abbildung 15: Aufforderung im Arbeitsheft zum Protokollieren



Abbildung 16: Anbringung und Unterteilung der Hilfsmittel in der Experimentalgruppe.

Die Experimentalgruppe erhielt zusätzlich materielle Unterstützungen in Form von vorgefertigten Lückentexten, Tabellen und Skizzen. Diese sollte als Hilfestellung beim Protokollieren dienen und sicherstellen, dass die Schüler ihre Ergebnisse während des selbstständigen Experimentierens aufschreiben. Für jedes Experiment

konnte zwischen den Repräsentationsformen Tabelle, Skizze oder Text gewählt werden (vgl.

Abbildung 16). Das Hilfsmittel konnte an der gekennzeichneten Stelle im Arbeitsheft eingeklebt werden. Die Interventionsmaßnahme der Gruppen BeginnN und BeginnM gestaltete sich durch die Reihenfolge im Besuch der jeweiligen Schülerlabore.

Pilotierung

Die Lerneinheit und der Wissenstest wurde vor Beginn der Vergleichsstudie mit einer 6. Klasse (N=28) des Kurfürst-Ruprecht-Gymnasiums in Neustadt an der Weinstraße pilotiert. Diese Klasse durchlief die gleiche Lerneinheit, Intervention, als auch Pre- und Post-Test, wie die beiden Gruppen der Hauptstudie. Bei der Überprüfung der Lerneinheit standen die Verständlichkeit der Arbeitshefte und die zeitliche Organisation der Experimente im Fokus. Durch die Beobachtung der Schüler während des Experimentierens und die anschließende Expertenbesprechung der Einheit mit dem betreuenden Lehrer wurden möglichst viele Fehlerquellen identifiziert und im Anschluss überarbeitet. Probleme im Stundenablauf und bei den Experimenten wurden auf Grundlage der Pilotierung bis zur Hauptstudie verbessert. Bei der Überprüfung der Testhefte und Rücksprache mit den Schülern wurde auf die Verständlichkeit der Fragen und die zeitliche Durchführung in 45 Minuten geachtet.

3.2.1 Methoden der Datenerhebung

Die verwendeten Daten dieser Arbeit stammen ausschließlich vom Schülerlabor Nawi-Werkstatt – die Daten vom Mathematiklabor werden von der Arbeitsgruppe Mathematikdidaktik Sekundarstufe genutzt. Die Datenerhebung erfolgte mittels Testhefte, Videoaufnahmen und Forscherhefte. Die Messzeitpunkte der Testhefte wurden entsprechend dem Design eines „Pre- / Post- / Follow-Up-Tests“ gewählt. Die Schüler wurden während der Testphasen und der Intervention über individuelle Identifikationsnummern gekennzeichnet. Die sechsstellige Nummer bestand aus vier schulinternen Ziffern plus zwei Ziffern des Geburtstages. Dies ermöglichte eine anonyme individuelle Rückverfolgung der Daten (TIEMANN & KÖRBS, 2014). Im Pre-Test wurden mittels Testhefte allgemeine demographische Daten (Alter, Geschlecht) und Migrationshintergrund der Probanden erhoben. Über alle drei Messzeitpunkte hinweg wurden das Fachwissen zu den durchgeführten Experimenten, die Motivation im Fach Naturwissenschaften ermittelt und Protokolle erstellt (vgl. Tabelle 5). Die Schule übermittelte die Noten des Jahreszeugnisses im Juli 2011 aller Schüler der Fächer Naturwissenschaften und Mathematik. Videoaufnahmen und Forscherhefte wurden während des Interventionszeitraums im Schülerlabor angefertigt.

Tabelle 5: Einsatz der Erhebungsinstrumente zu den jeweiligen Untersuchungszeitpunkten.

Zeitpunkt	Zu erfassende Variable	Erhebungsinstrument
Pre-Test	Demografische Daten Migrationshintergrund Motivation am Fach Fachwissen Experimente Protokolle	Abfrage im Testheft Abfrage im Testheft Items zur Motivation Entwickelter Fachtest Protokollieritem
Begleitend	Experimentierverhalten und Protokollieren Protokolle	Videoaufnahmen Forscherhefte
Post-Test	Motivation am Fach Fachwissen Experimente Protokolle	Items zur Motivation Entwickelter Fachtest Protokollieritem
Follow-Up-Test	Motivation am Fach Fachwissen Experimente Protokolle	Items zur Motivation Entwickelter Fachtest Protokollieritem
Anschließend	Experimentierverhalten und Protokollieren Einstellung zum Protokollieren	Interview

Die Forscherhefte über die drei Messzeitpunkte dienten als Instrument zur Erhebung der Protokolle. Diese stellen die abhängige Variable in Forschungsfrage 1.1 dar, welche auf Unterschiede bezüglich der unabhängigen Variablen untersucht wird. Durch die Videoaufnahmen und die Interviewdaten konnte die abhängige Variable in Forschungsfrage 1.2, der Prozess des Protokollierens, erhoben werden. Die Gruppenzugehörigkeit der Probanden wurde nicht zu den Messzeitpunkten erhoben, sondern über die Identifikationsnummer während der Intervention erfasst. Die weiteren Erhebungsinstrumente dienten zur Erfassung der hier verwendeten Kontrollvariablen (Demographische Daten, Migrationshintergrund). Weitere Kontrollvariablen sind die Noten im Fach Naturwissenschaften und im Fach Mathematik. Zusätzlich Instrumente (Lernwirksamkeit, Motivation am Fach, Video-Item) dienten zur Erhebung von Daten weiterer Projekte, welche diese Untersuchung flankierten. Deshalb sind die Konstrukte an dieser Stelle nicht genau aufgeführt, sondern dient nur der Vollständigkeit.

Forscherhefte

Jeder Schüler erhielt für den Interventionszeitraum ein Forscherheft, welches nach den drei Doppelstunden und den sieben Experimenten gegliedert war. Das Forscherheft wurde mit der jeweiligen Identifikationsnummer gekennzeichnet und den Schülern bei jeder Stunde wieder vorgelegt. Der vorliegende Datensatz der 180 Schüler gliedert sich in 90 Arbeitshefte mit je sieben Protokollen ohne Protokollierhilfe und 90 Arbeitshefte mit je sieben Protokollen mit Protokollierhilfe. Insgesamt fertigen die 180 Schüler im Verlauf der Doppelstunden zu jedem Experiment ein Protokoll an. Die 1260 Protokolle berechnen sich durch sieben Experimente von 180 Schülern. Davon weisen die Hälfte der Protokolle keine Protokollierhilfen auf. Krankheitsbedingt kam es in manchen Kleingruppen zu Ausfällen, die tatsächliche Protokollanzahl ist deshalb etwas geringer.

Videoaufnahmen

Videografiert wurde jeweils eine Kleingruppe der Experimentalgruppe aller sechs Klassen über die gesamte Zeit der drei Doppelstunden. Ziel war es, den Prozess des Experimentierens und des Protokollierens zu erfassen. Die Kamera Sony HDR-CX570 wurde mit einem Stativ möglichst unauffällig im Raum platziert. Im Fokus des Bildes stand der Gruppentisch, sodass alle Gruppenmitglieder im Bild waren. Das externe Bluetooth-Mikrofon Sony ECM-HW2 wurde für eine bessere Tonqualität zusätzlich montiert und der Empfänger verdeckt auf dem Gruppentisch gelegt. Die 1620 Minuten Bild- und Tonmaterial setzen sich aus sechs

Klassen von drei Doppelstunden à 90 Minuten zusammen. Aufgrund von technischen Problemen konnte nicht jede Doppelstunde aller Klassen gefilmt werden, weshalb die effektive Dauer der Videoaufnahmen kürzer ist.

Interview

Die Interviews wurden nach der Methode des problemzentrierten Interviews durchgeführt (n=10). Das problemzentrierte Interview ist eine Befragungsmethode der qualitativen Sozialforschung nach WITZEL (1982). Die Forschungsmethode der Befragung unterteilt sich in sieben Arbeitsschritte (BORTZ & DÖRING, 2006): 1. Inhaltliche Vorbereitung, 2. Organisatorische Vorbereitung, 3. Gesprächsbeginn, 4. Durchführung und Aufzeichnung des Interviews, 5. Gesprächsende, 6. Verabschiedung und 7. Gesprächsnotizen. Die Verwendung offener Fragestellungen ermöglicht dem Befragten den Gesprächsverlauf mitzugestalten, während der Interviewer durch gezieltes Nachfragen die Richtung steuert (BORTZ & DÖRING, 2006; WITZEL, 1982). Zu Beginn des problemzentrierten Interviews werden Sondierungsfragen gestellt, die die Atmosphäre auflockern und in das Thema einführen (BORTZ & DÖRING, 2006). Als Sondierungsfrage wurde nach dem aktuellen Thema im Fach Naturwissenschaften gefragt, um eine passende Überleitung zur Intervention zu schaffen. Anschließend wurden die zentralen Leitfragen eingebracht. Zunächst befragten wir die Schüler mit drei Leitfragen zu ihrer generellen Einstellung bezüglich des Protokollierens. Im Anschluss sollten die Schüler von dem Entstehen ihrer Protokolle während der Intervention berichten. Dazu wurde den Schülern ein ausgewähltes von ihnen verfasstes Protokoll vorgelegt. In Abbildung 17 ist der Interviewleitfaden zusammengefasst. Aufgezeichnet wurden die Interviews mit Hilfe eines Diktiergeräts.

- Wenn du in der Schule experimentierst, schreibst du etwas auf?
- Wie findest du das Aufschreiben?
- Was denkst du, warum schreibt man beim Experimentieren etwas auf, beispielsweise als Wissenschaftler oder in der Schule?
- Situationsbedingte Fragen zum eigenen Arbeitsheft
- Hattest du irgendwelche Probleme?
- Hast du alleine protokolliert oder mit deinem Gruppenpartner zusammen?

Abbildung 17: Leitfragen des problemzentrierten Interviews.

3.2.2 Konzeption der Lerneinheit

In dieser Untersuchung führten sechs Schulklassen der Jahrgangsstufe 6 drei Einheiten zum Thema „Stoffe im Alltag“ im Schülerlabor Nawi-Werkstatt durch. Die Konzeption der Einheit orientierte sich am Rahmenlehrplan des Fachs Naturwissenschaften für die 5. und 6. Klasse des Landes Rheinland-Pfalz. Der Rahmenlehrplan ist unterteilt in acht Themenfelder, wovon vier in Klassenstufe 5 und weitere vier Themenfelder in Klasse 6 unterrichtet werden (MINISTERIUM FÜR BILDUNG, WISSENSCHAFT, JUGEND UND KULTUR, 2010). Schüler sollen nach dem Lehrplan Stoffe durch selbst entwickelte Kriterien ordnen und Kriterien geleitet experimentelle Untersuchungen an Stoffen durchführen (ebd.). Das zu vermittelnde Fachwissen umfasst vor allem die spezifischen Stoffeigenschaften und die Möglichkeit Stoffe nach diesen Kategorien ordnen zu können. Dieses Projekt wurde, wie in Abschnitt „Design und Stichprobe“ bereits begründet, passend für das Themenfeld 7 konzipiert. Die Inhalte der Experimente sowie die Arbeitsweisen bauen aufeinander auf. Tabelle 6 gibt einen Überblick über die Experimente und deren Inhalt. Jede Doppelstunde begann mit einem gemeinsamen Auftakt, welcher das Thema der Stunde einleitete. Die Experimente waren nacheinander in der vorgegeben Reihenfolge zu bearbeiten. Die Kleingruppen bearbeiteten selbstständig die in Kisten vorbereiteten Experimente. Jede Kiste enthielt alle erforderlichen Materialien, Gerätelisten mit den Namen der Materialien sowie weitere Hilfestellungen zur Bearbeitung der Fragestellung. Nach dem Beenden eines Experimentes ging die Kleingruppe selbstständig zur Kiste für das nächste Experiment. Ungefähre Zeitvorgaben für jedes Experiment sollten von den Schülern eingehalten werden, um alle Inhalte der Doppelstunde bearbeiten zu können. Für schnelle Gruppe gab es in jeder Doppelstunde ein Zeitpuffer-Experiment. Dieses war inhaltlich auf die Stunde abgestimmt, jedoch inhaltlich nicht relevant für den Fachtest.

Tabelle 6: Übersicht der eingesetzten Experimente im Verlauf der Intervention.

Stunde	Experiment	Titel	Inhalt
1	1	Was sprudelt denn da?	Die Inhaltsstoffe der Brause auf ihre Eigenschaften überprüfen.
	2	Das Gas der Brause	Das entstehende Gas über seine Eigenschaften identifizieren.
2	3	Welches Gas ist im unbekanntem Luftballon?	Identifizieren eines unbekanntes Gases über den Vergleich der Löslichkeit bekannter Gase in Wasser.
	4	Entsteht überall gleich viel Kohlenstoffdioxid?	Vergleich von CO ₂ -Gehalt verschiedener Gasgemische
	5	Unbekannte Brausetablette	Vergleich von Brausetabletten mit Hilfe der unterschiedlichen Masse an Kohlenstoffdioxid in Wasser.
3	6	Treibhauseffekt durch verschiedene Gase	Mit Hilfe eines Modellexperimentes den Einfluss von Gasen auf den Treibhauseffekt identifizieren.
	7	Wie entstehen diese Gase?	Identifizieren der Reaktionen bei denen die Gase Kohlenstoffdioxid und Lachgas entstehen.

Eigenständiges Arbeiten

Ziel der Schülerlaboreinheiten war ein eigenständiges Arbeiten der Schüler, um die Lehrvariable ausschließen zu können. Damit der eigenständige Übergang zum nächsten Experiment funktioniert, wurde ein Feedbacksystem in die Einheit eingebaut. Dieses System war für die Interventionsgruppen gleich gestaltet.

Jede Kiste enthielt eine Checkliste, um die Reihenfolge und den Ablauf eines jeden Versuchs nochmal zu veranschaulichen (vgl. Abbildung 18). Damit es zu keinen Problemen beim Versuchsaufbau und der Versuchsdurchführung kam, sollten die Schüler nach dem Lesen der Anleitung eine Bilderreihe der Versuchsdurchführung in die richtige Reihenfolge bringen. Der Versuchsaufbau wurde in einzelnen Kärtchen mit Bildern gezeigt. Diese Aufgabe motivierte

die Schüler zum Lesen der Versuchsbeschreibung und visualisierte gleichzeitig den Aufbau des Versuchs. Zur Kontrolle des Versuchsergebnis war nach jedem Experiment eine Frage zu beantworten (Abbildung 19). Beantwortete die Schülergruppe die Frage richtig, wurde sie zu der Kiste mit dem nächsten Experiment geleitet. War die Antwort auf die Frage falsch, wurde die Gruppe zu einer falschen Kiste geleitet. Ziel war es, die Kiste zu bekommen, in deren Deckel die eigene Gruppennummer steht, zum Beispiel: „Reserviert für Gruppe 1“.

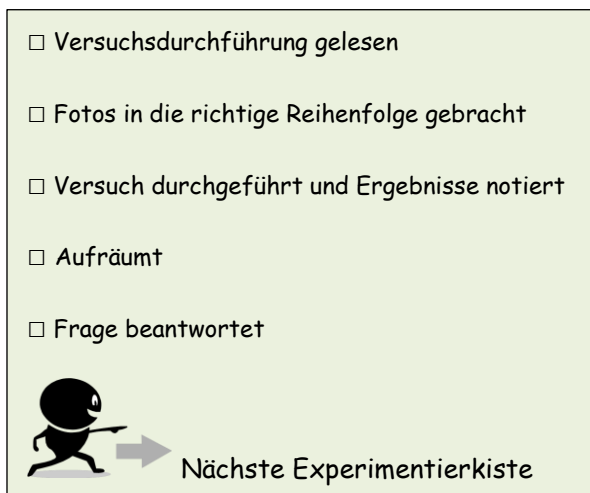


Abbildung 18: Checkliste als Hilfestellung für die Versuchsdurchführung.

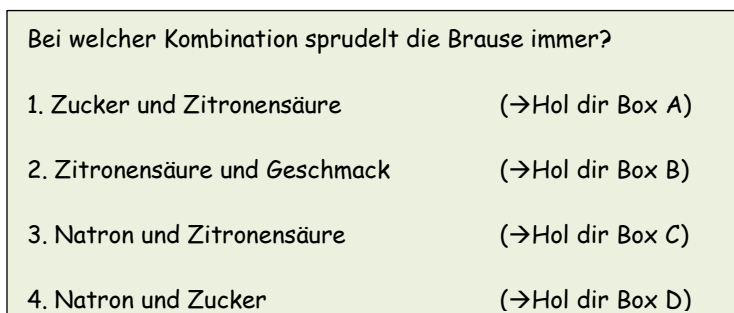


Abbildung 19: Beispiel für eine Kontrollfrage nach dem Experiment. Bei richtiger Antwort, kommt man zur richtigen Box.

Doppelstunde 1: „Was braust denn da?“

Zum Auftakt dieser Stunde wurde die Gasentwicklung beim Lösen einer Brausetablette in Wasser demonstriert. Dazu gab der Betreuer eine Brausetablette in ein Becherglas mit Wasser. Die Schüler beobachteten und hörten das Sprudeln des entstehenden Gases. Im Experiment „Was bringt die Brause zum Sprudeln?“ wurde zunächst Brausepulver unter dem Bionokular betrachtet. Ein Vergleich mit der Zutatenliste und dem Pulver zeigte, dass es sich bei dem Brausepulver um ein Stoffgemisch handelt. Die Kombination der einzelnen Zutaten unter dem Zusatz von Wasser sollte klären, welche Stoffe und welche Voraussetzungen für das Sprudeln der Brause verantwortlich sind. Ziel dieses Experiment war es, das die Schüler die Kombination von Natriumhydrogencarbonat und Zitronensäure mit Wasser als Voraussetzung für das Sprudeln der Brause identifizieren. Im zweiten Experiment „Welches Gas entsteht beim Sprudeln?“ galt es, das beim Sprudeln der Brause entstehende Gas zu identifizieren. Das Gas wurde dafür in einem Luftballon aufgefangen und auf seine Eigenschaften Dichte, Löschen von Feuer und Acidität überprüft. Die Ergebnisse der Experimente wurden mit den Steckbriefen der Gase Kohlenstoffdioxid, Distickstoffoxid² und dem Stoffgemisch Luft verglichen. Diese Steckbriefe enthielten die Eigenschaften der Gase beziehungsweise Gasgemische, welche vorab durch die Experimente überprüft wurden. Durch diese Vorgehensweise wurde Kohlenstoffdioxid als das bei der Brause entstehende Gas identifiziert.

Doppelstunde 2: „Wer braust wie stark?“

Das gemeinsame Üben der Technik des Einleitens eines Gases in einen Messzylinder durch Luftballone diente in der zweiten Doppelstunde als Auftakt-Experiment. Diese Technik fand in dieser Stunde eine zentrale Anwendung und es war nicht davon auszugehen, dass die Schüler damit bereits vertraut waren. Diese Übung erklärte den Schülern auch die Wasser-Verdrängung durch das Gas und das Ablesen der Werte. Im ersten Experiment der Stunde „Welches Gas ist im Luftballon?“ sollte ein unbekanntes Gas identifiziert werden. Durch das Lösen des Gases in Wasser und den Vergleich mit der Löslichkeit der Gase Kohlenstoffdioxid, Luft und Distickstoffoxid benannten die Schüler das Gas im Luftballon. Zusätzlich zu dem bereits gefüllten Luftballon mit dem unbekanntem Gas erhielten die Schüler Brausetabletten

² Der Begriff „Distickstoffoxid“ wurde als vereinfachter Begriff für den Stoff „Distickstoffmonooxid“ verwendet. Der Trivialname „Lachgas“ wurde aus didaktischen Gründen vermieden.

zum Herstellen von Kohlenstoffdioxid, eine Luftpumpe und Sahnekapseln zum Einfüllen von Distickstoffoxid. Für einen Vergleich der Gase musste jeweils das gleiche Volumen der Gase in Wasser gelöst werden. Dafür stand den Schülern ein Metallring zur Verfügung, der das Volumen des Luftballons vorgab. Ziel war es, dass die Schüler durch das Vergleichen der Ergebnisse Kohlenstoffdioxid als Gas aus dem unbekanntem Luftballon identifizieren. In einem zweiten Experiment „Entsteht überall gleich viel Kohlenstoffdioxid?“ wurde der Kohlenstoffdioxid-Gehalt verschiedener Gase verglichen. Die Schüler füllten mehrere Luftballons mit unterschiedlichen kohlenstoffdioxidhaltigen Gasen und leiteten diese anschließend in eine Bromthymolblau-Lösung ein. Zum Vergleich dienten die Gasgemische (1) welches entsteht, wenn Brause in Wasser gelöst wird, (2) Backpulver und Essig gemischt werden, sowie (3) Atemluft. Für einen Vergleich war es wichtig, die Gase jeweils nur so lange einzuleiten bis der Indikator von blau zu gelb umschlägt. Anschließend wurde das Volumen der Luftballons mit Metallringen gemessen. Das Gas im Luftballon mit dem kleinsten Metallring hatte die geringste Konzentration an Kohlenstoffdioxid, da mehr Gas eingeleitet werden muss, bis der Indikator umschlägt. Ziel dieses Experimentes war es, dass die Schüler unterschiedliche Gasgemische mit Kohlenstoffdioxid kennen lernen und diese aufgrund ihrer Konzentration unterscheiden können. Im dritten Experiment wurde eine „Unbekannte Brausetablette“ untersucht. Die Schüler bestimmten das Volumen des entstandenen Kohlenstoffdioxids in Wasser von einer unbekanntem und drei bekannten Brausetabletten. Durch einen Vergleich der Messwerte konnte die unbekanntem Brausetablette mit dem Namen einer bekannten Tablette benannt werden. In diesem Experiment stand die Methode des Messens von Gasvolumina im Vordergrund.

Doppelstunde 3: „Wo braust es überall?“

Zu Beginn der dritten Doppelstunde wurde ein Video über den Einfluss der Gase in der Atmosphäre auf den Treibhauseffekt gezeigt. Die Schüler widmeten sich im Experiment „Treibhauseffekt durch verschiedene Gase“ der Untersuchung treibhausrelevanter Gase und deren Einwirkung auf die Temperatur der Atmosphäre. Durch ein Modellexperiment mit den Gasen Luft, Kohlenstoffdioxid und Distickstoffoxid wurde der unterschiedliche Temperaturanstieg der Atmosphäre über eine gleiche Belichtungszeit mit einem Halogenscheinwerfer veranschaulicht. In dem Experiment „Wie entstehen diese Gase?“ wurde anschließend die Entstehung beziehungsweise die Zusammensetzung dieser drei Gase veranschaulicht. Die Entste-

Die Entstehung von Kohlenstoffdioxid wurde durch eine Verbrennung von Papier und das anschließende Einleiten des aufgefangenen Gases in Bromthymolblau veranschaulicht. Die Schüler wussten zu diesem Zeitpunkt bereits, dass Bromthymol in Anwesenheit von Kohlenstoffdioxid von blau zu gelb umschlägt. Die Entstehung des Gases Distickstoffoxid durch Düngung wurde über ein Bandolino-Spiel erklärt. Ein weiteres Bandolino erklärte die Zusammensetzung der Luft. Ziel dieser Station war es, die Entstehung von Kohlenstoffdioxid durch Verbrennung und Distickstoffoxid durch Düngung zu veranschaulichen.

3.2.3 Methoden der Datenauswertung

Zur Überprüfung der Unterschiede in den Protokollen durch die design- und interventionsbedingten Kriterien (Forschungsfrage 1.1), wurden zunächst die Schülerprotokolle nach entwickelten Kriterien ausgewertet. Die Ergebnisse dieser Bewertung wurden auf Unterschiede in den designbedingten (Experiment 1-7), interventionsbedingten (BeginnN-BeginnM, KG-EG) sowie weitere unabhängige Merkmalen (Nawi-Note) geprüft. Um Unterschiede im Prozess des Protokollierens zu veranschaulichen (Forschungsfrage 1.2), wurde dieser mit der Auswertung der Videoaufnahmen und der Interviews untersucht. Die gewonnenen Videodaten wurden quantifiziert und auf die Beeinflussung durch interventionsbedingte Merkmale getestet. Die Schülersaussagen aus den Interviews wurden ebenso quantifiziert und Ankerbeispiele ausgewählt. Schüler, die nicht zu allen drei Interventionszeitpunkten anwesend waren, trugen zur weiteren Auswertung nicht bei. Deren Forscherhefte und Testhefte trugen zusätzlich zur Entwicklung der Auswertungsmethoden bei.

Forscherhefte

Die Protokolle der Forscherhefte wurden mittels qualitativer Inhaltsanalyse ausgewertet, anschließend quantifiziert und die Variablen auf Unterschiede überprüft.

Die Entwicklung des Kategoriensystems wurde mittels der Forscherhefte von je zwei Klassen der zwei Schulen mit je sieben Experimenten entwickelt. Die Protokolle gehörten zu Teilnehmern sowohl der Gruppe Beginn Nawi-Werkstatt (BeginnN) oder Beginn Mathematik-Labor (BeginnM), als auch der Untersuchungsgruppe mit Protokollierhilfen (EG) oder ohne Protokollierhilfen (KG) an. Die qualitative Inhaltsanalyse ist ein Standardverfahren der qualitativen Sozialforschung mit dem Ziel methodisch kontrolliert und intersubjektiv überprüfbar die Inhalte von Medien zu analysieren (MAYRING, 2002; BORTZ & DÖRING, 2006). Hier wurde das strukturierende dreistufige Verfahren der qualitativen Inhaltsanalyse nach MAYRING verwendet. Zunächst wurden deduktiv Kategorien entwickelt, denen die zu untersuchenden Daten in der zweiten Phase zugeordnet wurden (MAYRING, 2002). Im Anschluss erfolgte die induktive Rückkopplung zur Überarbeitung der gewonnenen Kategorien (ebd.). Aus den Protokollinhalten wurden induktiv neue, vorher nicht bedachte, Kategorien herausgearbeitet oder zu übergreifenden Kategorien zusammengefasst (ebd.). Für die Bewertung der Protokolle ergeben sich drei Kategorien, welche in Tabelle 7 benannt und definiert sind. Die Inhal-

te dieser Kategorien sind in jedem Protokoll vorhanden und können dadurch zuverlässig ausgewertet werden.

Tabelle 7: Kategorien für die Bewertung aller Forscherhefte.

Kategorie	Definition
Wortanzahl	<i>Wie viele Wörter wurden in einem Protokoll geschrieben?</i>
Verwendung des Protokollierbereiches	<i>Wurde der vorgegebene Protokollierbereich verwendet?</i>
Repräsentationsform	<i>Welche Repräsentationsform der Protokolle wurde verwendet?</i>

Die Forscherhefte mit jeweils sieben Experimenten der anderen zwei Klassen der zwei Schulen, wurden nach der Entwicklung auf die Ausprägung der Kategorien des Kategoriensystems bewertet. Der Kodierplan ist im Anhang (Anhang II) aufgeführt und erhöht die Objektivität der Datenauswertung. Die Ausprägungen wurden in MICROSOFT EXCEL 2010 eingegeben und deskriptiv ausgewertet. Die Protokolle der Probanden gehörten wiederum der Gruppe Beginn Nawi-Werkstatt (BeginnN) oder Beginn Mathematik-Labor (BeginnM), als auch der Untersuchungsgruppe mit Protokollierhilfen (EG) oder ohne Protokollierhilfen (KG) an. Durch diese Auswahl wurden Störvariablen, wie die Lehrperson oder Schulzugehörigkeit ausgeschlossen und trotzdem ist eine ausreichend große Stichprobe an Protokollen mit $n=441$ für die Entwicklung und die Auswertung gegeben.

Anschließend wurde eine Auswahl von zwei vergleichbaren Protokollen (Experiment 3 und 4) der Stichprobe auf Unterschiede in den Kategorien zu den unabhängigen Variablen geprüft. Die Stichprobe der Protokolle für diese Berechnungen umfasste $n=126$ und berechnet sich aus vier Klassen für jeweils nur zwei Experimente mit je drei bis vier Schülern pro Kleingruppe. Davon gehörten 31 Schüler zur Kontrollgruppe und 32 Schüler zur Experimentalgruppe. Die Kategorien der Protokollbewertung wurden auf Unterschiede auf die Variablen Gruppenzugehörigkeit Beginn Nawi-Werkstatt (BeginnN) oder Beginn Mathe-Labor (BeginnM), Zugehörigkeit zur Kontrollgruppe oder Experimentalgruppe, sowie Nawi-Schulnote (sehr gut – ungenügend) geprüft. Mit Hilfe des Statistikprogramms R.STUDIO VERSION 0.98.1091 wurde ein Vergleich der Mittelwerte (χ) und der Standardfehler (σ) der Ausprägung von unabhängigen Variablen und den abhängigen Variablen der Kategorien der Protokollbewertung heran-

gezogen. Eine Unterschiedsprüfung und Berechnungen der Fehlerwahrscheinlichkeit p wurde je nach Datenlage mit dem Testverfahren der ANOVA ohne Messwiederholung oder einem t-Test für unabhängige Stichproben durchgeführt. Das Signifikanzniveau α des p -Wertes wird nach FISCHER (1935) für alle weiteren Verfahren wie folgt festgelegt: $p < 0.05 =$ signifikant *, $p < 0.01 =$ hochsignifikant **, $p < 0.001 =$ höchstsignifikant *** (EID, GOLLWITZER & SCHMITT, 2011). Bei einem signifikanten Ergebnis eines multiplen Paarvergleiches wurden alle Mittelwerte paarweise mit dem Post-hoc-Verfahren mit einer α -Fehler-Korrektur (nach Bonferroni) durchgeführt. Zur Berechnung der Effektgröße der Varianzanalysen wurde die Effektgröße f nach Cohen gewählt. Die Effektgröße f nach Cohen wird wie folgt zugeteilt: $f \approx 0.14 =$ klein, $f \approx 0.35 =$ mittel, $f \approx 0.57 =$ groß (EID ET AL., 2011). Als Effektgröße für den t-Test wurde die standardisierte Berechnung nach Cohens d verwendet. Die Interpretation der Ergebnisse von Cohens d wird wie folgt beansprucht: $0.2-0.4 =$ kleiner Effekt, $0.5-0.7 =$ mittlerer Effekt, $>0.7 =$ großer Effekt (EID ET AL., 2011).

Als Zusammenhangsmaß von Werten bestimmen zu können wurde die Produkt-Moment-Korrelation nach PEARSON gewählt. Je größer der Wert, umso stärker ist der Zusammenhang. Nach COHEN (1988) lassen sich die Korrelationen der Produkt-Moment-Korrelation wie folgt interpretieren: $r_{XY} \approx 0,10$ entspricht einem schwachen Zusammenhang, $r_{XY} \approx 0,30$ einem mittleren Zusammenhang und $r_{XY} \approx 0,50$ einem starken Zusammenhang (EID ET AL., 2011).

Videoaufnahmen

Für die Videoaufzeichnungen wurden ebenfalls mittels qualitativer Inhaltsanalyse Kategorien gebildet, anschließend quantifiziert und die entstandenen Variablen auf Unterschiede geprüft.

Die Videoanalyse umfasst je eine Kleingruppe der Experimentalgruppe aller sechs Klassen. Insgesamt lagen von sechs Kleingruppen Videoaufzeichnungen von 42 Experimenten vor. Die Stichprobengröße $n=42$ berechnet sich aus sechs Kleingruppen über sieben Experimente. Aufgrund technischer Probleme und Vergleichbarkeit der Experimente wurden tatsächlich jedoch nur $n=12$ Aufnahmen ausgewertet und überprüft (sechs Kleingruppen der vergleichbaren Experimente drei und vier). Wie auch im Forscherheft wurden die drei Phasen der qualitativen Inhaltsanalyse nach MAYRING (2002) umgesetzt. Es entstanden induktiv fünf Kategorien die in Bezug auf den Prozess des Protokollierens über alle Videoaufnahmen hinweg

beobachtet werden konnten. Die verwendeten Kategorien sind in Tabelle 8 und der entwickelte Kodierplan ist im Anhang (Anhang III) aufgeführt.

Tabelle 8: Kategorien für die Bewertung der Videoaufnahmen.

Definition	
Protokollieren	<i>Protokollieren die Schüler?</i>
Sozialform Protokollieren	<i>In welcher Sozialform protokollieren die Schüler?</i>
Experimentieren	<i>Experimentieren die Schüler?</i>
Sozialform Experimentieren	<i>In welcher Sozialform experimentieren die Schüler?</i>
Absprache Protokollierhilfe	<i>Wurde die Protokollierhilfe in Absprache ausgesucht?</i>

In der zweiten Phase erfolgte eine zeitabhängige Auswertung der Videoaufnahmen (BRÜCKMANN & DUIT, 2014). Das bedeutet, dass immer ein Intervall kodiert wurde und das darauf folgende Intervall nicht (ebd.). Durch diese Art der Kodierung sollte gewährleistet werden, dass der Prozess veranschaulicht wird und sich die Auswertung nicht nur auf ein bestimmtes Ereignis beschränkt. Die Messung begann, sobald ein Kleingruppenmitglied die Kiste mit den Experimentiermaterialien für das jeweilige Experiment auf den Tisch stellte. Nun wurde über die Dauer eines Experiments im Abstand von zwei Minuten jeweils für 20 Sekunden Videomaterial eingesehen. In dieser Sequenz wurden alle Aktivitäten oder Äußerungen der gesamten Kleingruppe mit dem Kategoriensystem bewertet. Die Messung endete nach einer zweifach ausbleibenden Experimentier- oder Protokolliertätigkeit der gesamten Kleingruppe. Dadurch war das Messzeitpunktraster genau definiert und die Datenerhebung von Experiment zu Experiment vergleichbar. Danach erfolgte die Quantifizierung der Merkmalsausprägungen mit absoluten und relativen Häufigkeiten mit MICROSOFT OFFICE EXCEL.

Interviews

Die aufgenommenen Interviews (n=10) wurden transkribiert, mittels qualitativer Inhaltsanalyse ausgewertet, mit Ankerbeispielen belegt und anschließend quantifiziert.

Die Interviews wurden in ein Gesamttranskript verschriftlicht. Das Transkript liegt im Dialekt der Schüler vor und wurde nicht auf sprachliche Fehler bereinigt. Pausen sind mit „...“ markiert und mimische und gestische Kommunikation ist in „*___*“ aufgeführt. Die Schülernamen sind auf den Anfangsbuchstaben reduziert, um die Anonymität zu wahren und trotzdem

zu erkennen von wem welche Aussage stammt. Die transkribierten Interviews wurden nach den drei Phasen der qualitativen Inhaltsanalyse nach MAYRING (2002) ausgewertet (vgl. Forscherhefte). Der entwickelte Kodierplan ist im Anhang (Anhang IV) zu finden. Um im weiteren Verlauf der Datenauswertung Textstellen leichter zu analysieren und zuzuordnen, wurde mit Ankerbeispielen gearbeitet (ebd.). Ankerbeispiele sind typische Textstellen, die exemplarisch für eine Kategorie im Kodierplan stehen (ebd.). Diese Ankerbeispiele dienten als vergleichende Belege, um die Ergebnisse der Auswertung zu untermauern. Die Quantifizierung erfolgte mit MICROSOFT OFFICE EXCEL 2010, vergleichbar mit der Vorgehensweise Forscherheftanalyse.

3.2.4 Untersuchungsergebnisse

Die Untersuchungsergebnisse und deren anschließende Diskussion sind bezüglich der Forschungsfragen so gegliedert, dass die design- und interventionsbedingten Unterschiede von Protokollen und des Prozess des Protokollierens gemeinsam angesprochen werden. Die Darstellung der Ergebnisse beschränkt sich auf einen ausgewählten Teil der möglichen Vergleiche zum Beispiel in den Kategorien der Protokollauswertung. Forschungsfrage 1.1 wird mit den Daten der Protokollauswertung durch einen Vergleich der Experimente, der Untersuchungsgruppen BeginnM und BeginnN sowie Kontrollvariablen der Untersuchung zum Beispiel Nawi-Note beantwortet. Ankerbeispiele der Interviewdaten unterlegen die Ergebnisse. Forschungsfrage 1.2 wird mit den Daten der Video- und Interviewauswertung entgegnet.

Abbildung 20 zeigt die gemittelte Wortanzahl der Protokolle im Vergleich über die sieben Experimente. Die Wortanzahl der Protokolle schwankt von Experiment zu Experiment von gemittelt zehn bis 34 Wörtern. Lediglich Experiment drei und vier wurden im Schnitt mit gleich vielen Wörtern protokolliert.

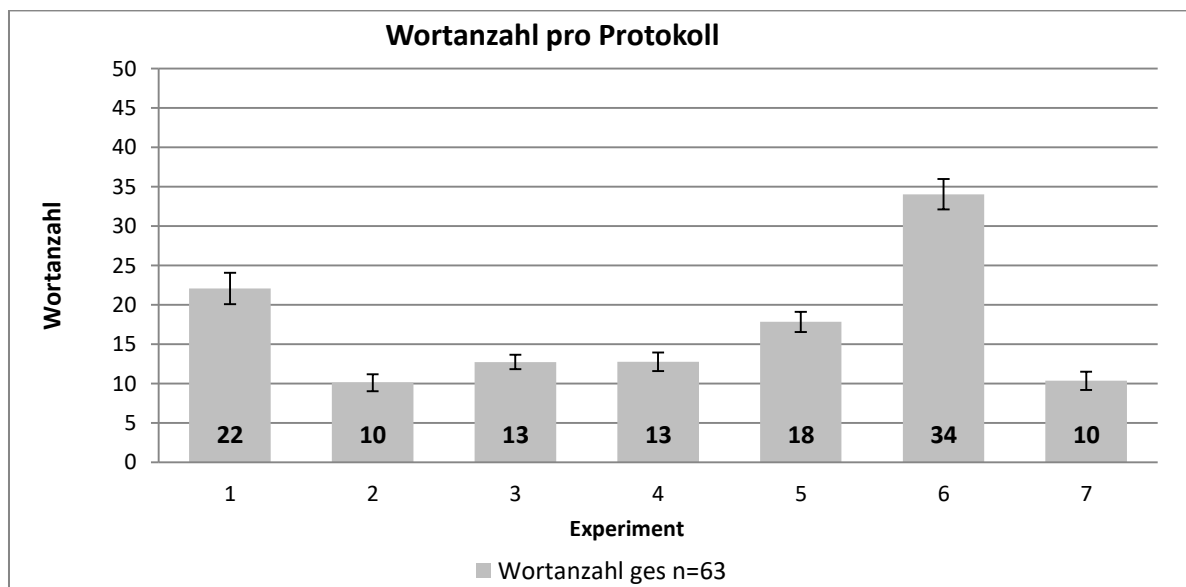


Abbildung 20: Gemittelte Wortanzahl der Protokolle mit Standardfehler im Vergleich der sieben Experimente. Die Stichprobengröße $n = 63$ berechnet sich aus 441 Protokollen geteilt durch 7 Experimente.

Abbildung 21 zeigt die gemittelte Wortanzahl der Protokolle zu Experiment 3 und 4 im Vergleich von Kontroll- und Experimentalgruppe. Diese Stichprobe der Protokollauswertung umfasst aufgrund der Vergleichbarkeit nicht alle Probanden aus Abbildung 20. Die Wortanzahl der Experimentalgruppe liegt bei $\chi = 10,0$ und $\sigma = 0,8$ in lila, die Wortanzahl der Kontrollgruppe mit $\chi = 15,5$ und $\sigma = 1,2$ in beige. Die Unterschiedsprüfung mit dem t-Test ergab eine

signifikante Fehlerwahrscheinlichkeit und einen mittleren Effekt ($t(98.5)=-3.97$, $p=0.0001^{***}$, $d=0.69$).

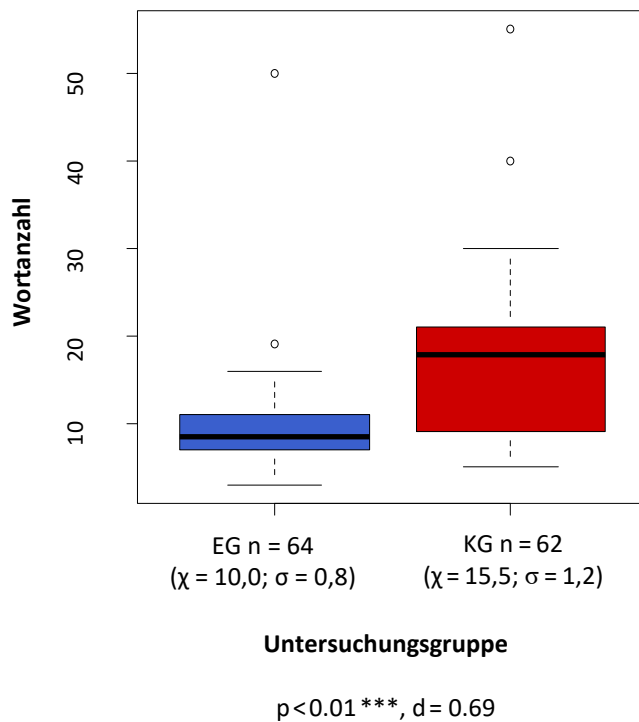


Abbildung 21: Wortanzahl der Protokolle aus den Untersuchungsgruppen EG und KG.

Die unterschiedliche Wortanzahl der Untersuchungsgruppen kann durch die Wahl der Repräsentationsform eines Protokolls erklärt werden und wird in Abbildung 22 veranschaulicht. Die gemittelte Wortanzahl der Repräsentationsform Text (grün) liegt bei $\chi=17,9$ und $\sigma=1,8$, der Repräsentationsform Tabelle (gelb) bei $\chi=10,4$ und $\sigma=0,6$ und der Repräsentationsform Skizze (braun) mit $\chi=16,0$ und $\sigma=3,0$. Die Unterschiedsprüfung hat eine Fehlerwahrscheinlichkeit der ANOVA ergab ein signifikantes Ergebnis und einen mittleren Effekt ($F(118)=20.21$, $p < 0.01^{***}$, $f=0.42$). Das Post-hoc-Verfahren ergab einen signifikanten Unterschied der Wortanzahl von Tabelle und Text ($p < 0.01^{***}$).

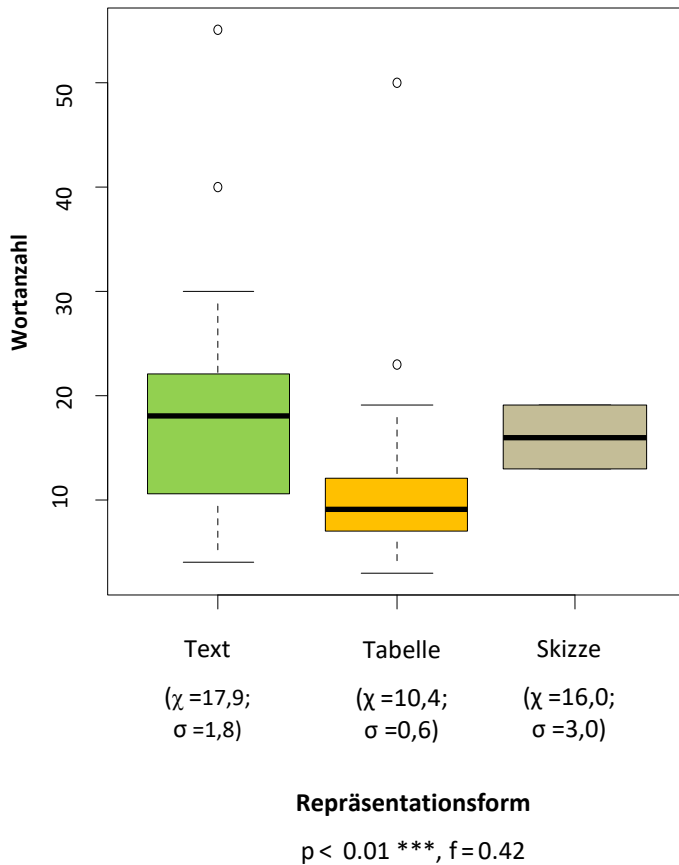


Abbildung 22: Wortanzahl der Protokolle in Abhängigkeit der Repräsentationsform der Protokolle.

Wie Abbildung 23 veranschaulicht, verwendeten die Schüler der Untersuchungsgruppe KG mehr als doppelt so häufig die Repräsentationsform Text als die Probanden der Untersuchungsgruppe EG. In beiden Untersuchungsgruppen dominierte die Repräsentationsform Tabelle. Da die Probanden der Untersuchungsgruppe EG zu 35% häufiger Tabellen nutzten, liegt die Wortanzahl der Protokolle in Abbildung 23 deutlich unter der Wortanzahl der Protokolle aus der Untersuchungsgruppe KG. Die Repräsentationsform Skizze wurde dagegen kaum genutzt und trägt deshalb nicht zur Interpretation bei.

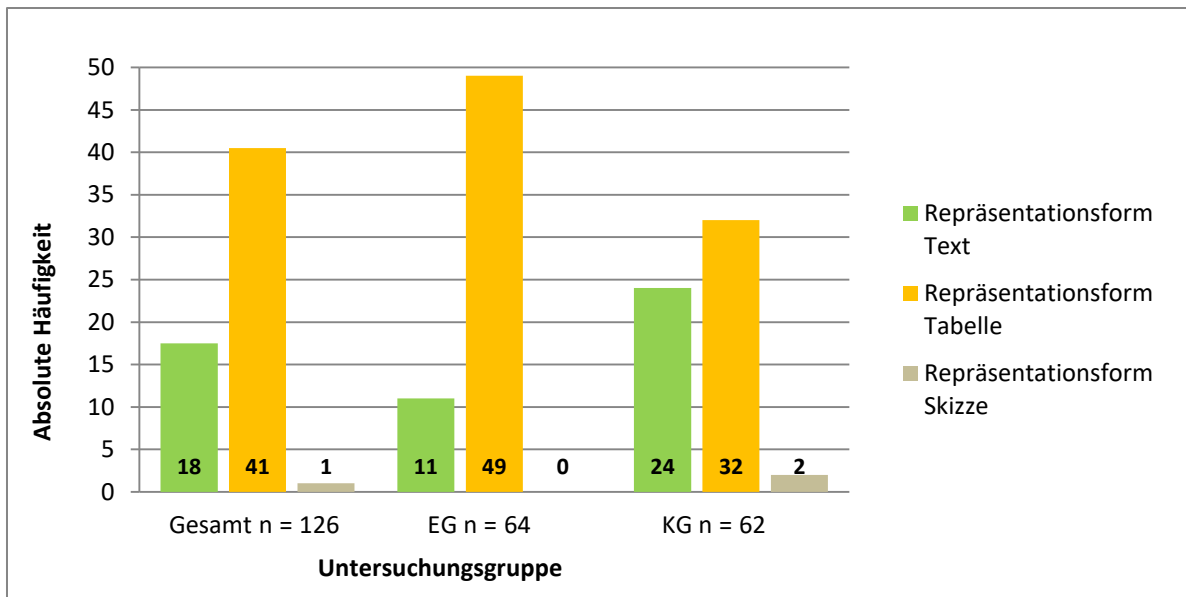


Abbildung 23: Absolute Häufigkeit der gewählten Repräsentationsform in Abhängigkeit der Untersuchungsgruppe EG und KG mit dem Durchschnitt beider Untersuchungsgruppen als Vergleich.

Im verwendeten Arbeitsheft war eine Blankoseite als Protokollierbereich vorgegeben. Abbildung 24 zeigt, wie häufig dieser Protokollierbereich pro Experiment verwendet wurde. In der Regel nutzten die Probanden ausschließlich den zur Verfügung stehenden Protokollierbereich, allerdings weichen Experiment 2 und 6 davon ab.

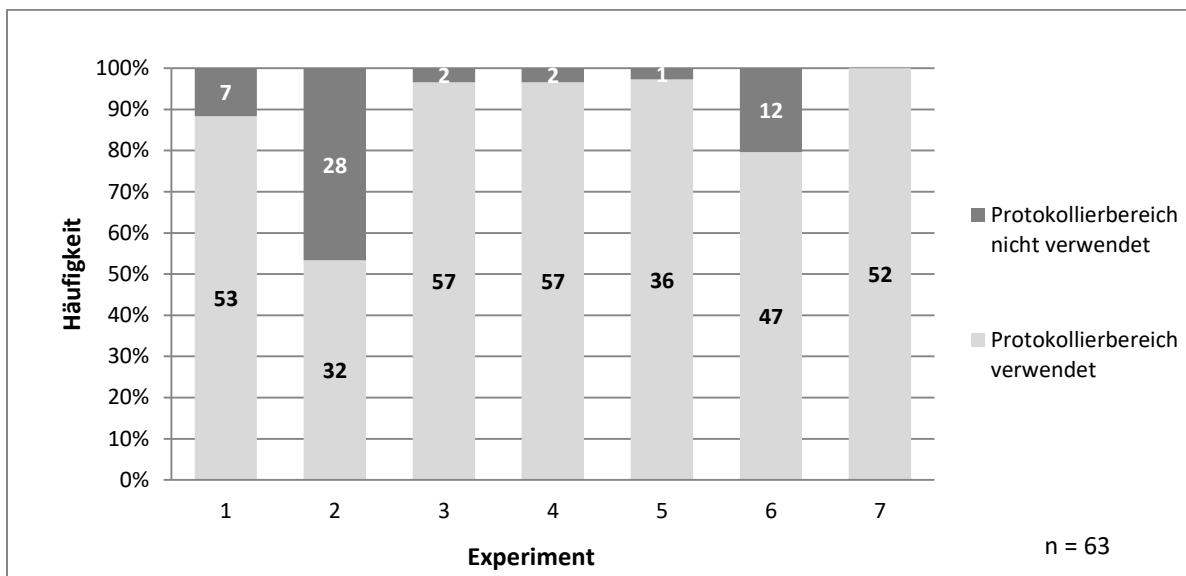


Abbildung 24: Relative und absolute Häufigkeit der Verwendung des Protokollierbereichs pro Experiment. Ausschließliche Verwendung des vorgegebenen Protokollierbereichs in hellgrau, abweichende Notizen vom vorgegebenen Protokollierbereich in dunkelgrau.

Bei den Arbeitsheften der Experimente 2 und 6 befand sich der Protokollierbereich auf der Rückseite. Die Probanden des Interviews bestätigen die Ergebnisse der Abbildung:

J.: „Also des Bessere war, wo auf der gleichen Seite das auch war. Ähm,...weil da eben nicht umblättern musst und erst die Frage lesen musst.“

*T.: „Ja, des hab' ich ja gar nicht gefunden *lacht*. Hab' ich gar nicht geseh'n.“*

Die Abweichung zur angesetzten Stichprobengröße von $n=63$ geht aus den absoluten Häufigkeiten der einzelnen Experimente aus Abbildung 24 hervor und beträgt bis zu 26 %, beispielsweise in Experiment 5. Manche Kleingruppen gerieten gegen Ende der Doppelstunde in Zeitnot, weshalb nicht alle Probanden alle Experimente durchgeführten oder protokollierten. Auch die Interviewdaten belegen dies:

M.: „Mhmm. Habs nicht mehr ganz geschafft, weil die Zeit nicht mehr so doll gereicht hat.“

Da sich sowohl der Arbeitsauftrag als auch die Gestaltung des Arbeitshefts von Experiment zu Experiment unterschieden und die Experimente innerhalb einer Doppelstunde in verschiedenen zeitlichen Phasen stattfanden, wird für die Ergebnisse im Folgenden lediglich auf die durch Abbildung 20 und Abbildung 24 veranschaulichten vergleichbaren Experimente 3 und 4 zurückgegriffen. Dadurch sind nur diese Experimente untereinander vergleichbar, um weitere Unterschiede prüfen zu können. So ergeben sich die geringeren Stichproben in den folgenden Abbildungen.

Die Probanden des Interviews gaben bei 85 % der Äußerungen auf die Frage „Hattest du irgendwelche Probleme?“ Probleme an, die sich auf das Protokollieren bezogen. Die 14 Äußerungen zu sonstigen Problemen lassen sich grob in die Bereiche Anforderungsniveau, Lernfluss und Motivation gliedern.

Anforderungsniveau:

V.: „Früher in der 5. Klasse mussten wir halt den Stoff halt genau gucken und halt das sind dann immer so ganz kleine Teile und das ist halt schwierig das genau abzumalen.“

Lernfluss:

M.: „Ja aber, das war so. Da wurden wir beim Schreiben unterbrochen, da mussten wir hier dann weiter machen.“

Motivation:

J.: „...weil ich andauernd Stift rausholen und...manchmal tut mir auch mein Knöchel weh.“

J.: „...Also, das was ich heute aufgeschrieben hätte, das wär' nich so gut geworden. Weil ich ja nicht so ausführlich geschrieben hab'.“

F.: „Da war ich zu faul zum Schreiben.“

Weitere Angaben zu den Problemen handelten von dem Umgang mit der vorgegebenen Protokollierhilfe, besonders wenn diese ausgefüllt werden sollten:

F.: „Mhmm. Des da hab' ich nicht so ganz kapiert, mit dem, wie wir die...Werte aufschreiben sollen...Ja, ob man da die Zahlen hinschreiben soll oder die Farben...“

M.: „Bei hier bei dem Wort hat das nicht so genau hin gepasst.“

Unabhängig von der Repräsentationsform des Protokolls oder ob Protokollierhilfen zur Verfügung stehen, schrieb die Gruppe BeginnM circa ein Viertel mehr Wörter als BeginnN (vgl. Abbildung 25). Die Wortanzahl BeginnM hat einen Mittelwert $\chi=14,5$ und $\sigma=1,1$ in rot, Wortanzahl BeginnN mit $\chi=11,0$ und $\sigma=0,9$ in blau. Die Fehlerwahrscheinlichkeit des t-Tests ergab ein signifikantes Ergebnis und einen kleinen Effekt ($t(118)=2.39$, $p=0.02^*$, $d=0.43$).

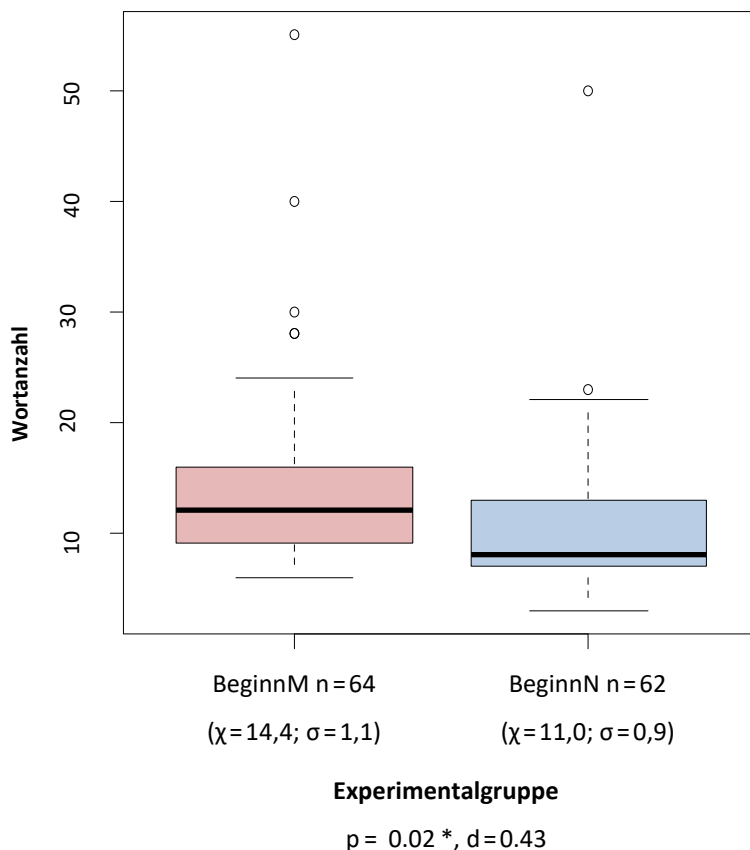


Abbildung 25: Wortanzahl der Protokolle der Experimente drei und vier aus den Gruppen BeginnN und BeginnM.

Neben dieser unabhängigen Variablen des Untersuchungsdesigns wurden weitere Unterschiede in der Wortanzahl gefunden. Die Schulnote der Probanden im Fach Naturwissenschaften beeinflusst signifikant die Anzahl an geschriebenen Wörtern in einem Protokoll. Die Annahme, je besser die Nawi-Schulnote desto ausführlicher das Protokoll, wurde durch die Datenauswertung falsifiziert und in entgegengesetzter Richtung bestätigt. Abbildung 26 veranschaulicht, je besser die Nawi-Schulnote, desto kürzer ist das angefertigte Protokoll. Schüler mit einer sehr guten Nawi-Schulnote schrieben zu rund 40% weniger als Schüler mit einer ausreichenden Nawi-Schulnote. Schüler mit der Note sehr gut benutzten $\chi=9,6$ und $\sigma=0,6$, Schüler mit der Note gut $\chi=13,2$ und $\sigma=1,0$ Wörter, Schüler mit der Note befriedigend $\chi=13,2$ und $\sigma=2,4$ und Schüler mit der Note ausreichend mit $\chi=16,2$ und $\sigma=3,3$. Die Unterschiedsprüfung der ANOVA ergab eine signifikante Fehlerwahrscheinlichkeit jedoch keinen nachweisbaren Effekt ($F(104)=5.01$, $p=0.03^*$, $f=0.07$). Die Regressionsgerade in Abbildung 26 zeigt einen schwache, positive Korrelation ($r=0.2$) zwischen der Nawi-Schulnote und der Wortanzahl.

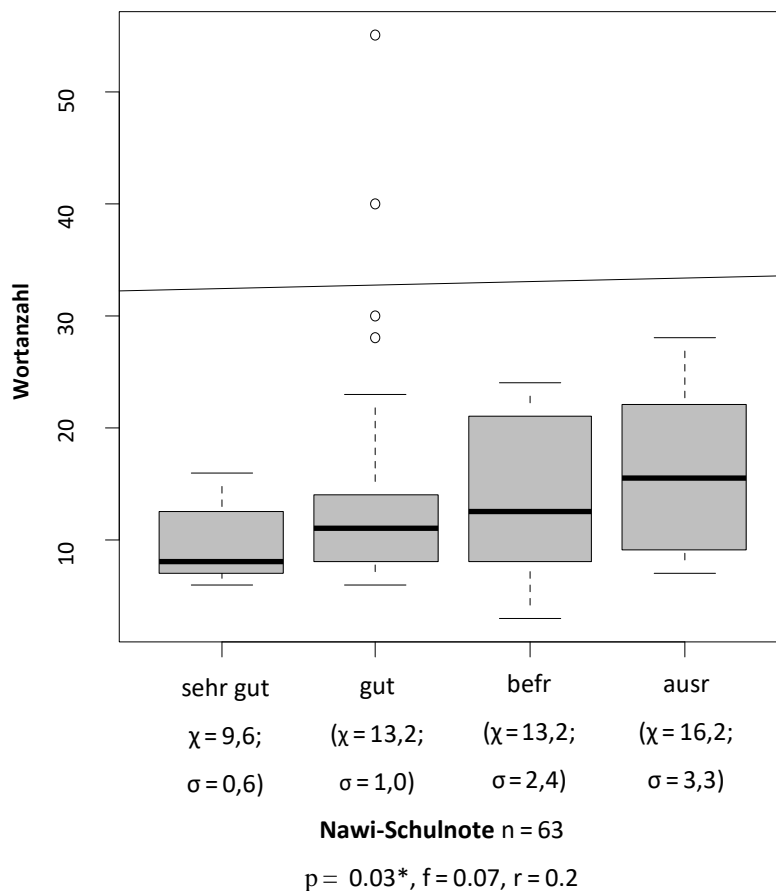


Abbildung 26: Wortanzahl der Protokolle aus Experiment drei und vier in Bezug auf die Nawi Schulnote. Gerade als lineare Regression von den Noten sehr gut bis ausreichend.

Bei den Interviews gaben 9 von 10 der Befragten an, aus der Schule Vorkenntnisse mit Versuchsprotokollen mitzubringen, obwohl nach Angabe der Lehrkräfte das Protokoll im Unterricht noch nicht thematisiert wurde. Es zeigte sich jedoch eine unterschiedliche Form der Vorerfahrung. Das Niveau der Vorerfahrung zum Protokollieren schwankt bei den Befragten. Einige Schüler kannten den genauen Aufbau eines Protokolls:

T.: „Also als erstes schreibt man halt hin, was man überhaupt macht. Dann, was...man dafür braucht und dann die Durchführung, Beobachtung und's Ergebnis.“

Andere verstanden unter einem Protokoll eher einen Merksatz:

*J.: „Ja, dann schreiben wir so Texte auf und ja, auf die Tafel und dann haben, schreiben wir's so auf'n Blatt und wenn's dann richtig ist, machen wir *kichert* nen roten Kasten drum.“*

Neben dem „Datum“ brachten die Schüler ein Protokoll auch mit Fachbegriffen in Verbindung:

J.: „Durchführung, Beobachtung“ und alles...und da müssen wir alles schreiben...Chemikalien...Da müssen wir das schon ausführlich schreiben. Und eben auch mit Fachbegriffen.“

Die Schüler wissen, welche Arbeitsschritte in einem Protokoll verlangt sind und kennen deren Relevanz.

Durch die Videoaufnahmen wurde die Sozialform beim Protokollieren und Experimentieren, die Wahl der Protokollierhilfe sowie Häufigkeit dieser Tätigkeiten aufgenommen (n=12). Die Frequenz der Tätigkeiten unterschied sich je nach Gruppenzugehörigkeit zu BeginnM oder BeginnN. Die gemittelten Tätigkeiten für die zwei ähnlichen Versuche sind in Abbildung 27 dargestellt. Auf der Skala entspricht „0“ keiner Tätigkeit und „1“ einer Tätigkeit, entweder Protokollieren oder Experimentieren. Innerhalb eines Versuchs protokollierten die Probanden einer Kleingruppe im Durchschnitt rund fünfmal weniger als sie experimentierten. Dieses Verhältnis änderte sich im Vergleich der Gruppen BeginnM und BeginnN. BeginnM protokollierte bei gleich bleibender Experimentiertätigkeit zu 55 % häufiger als BeginnN.

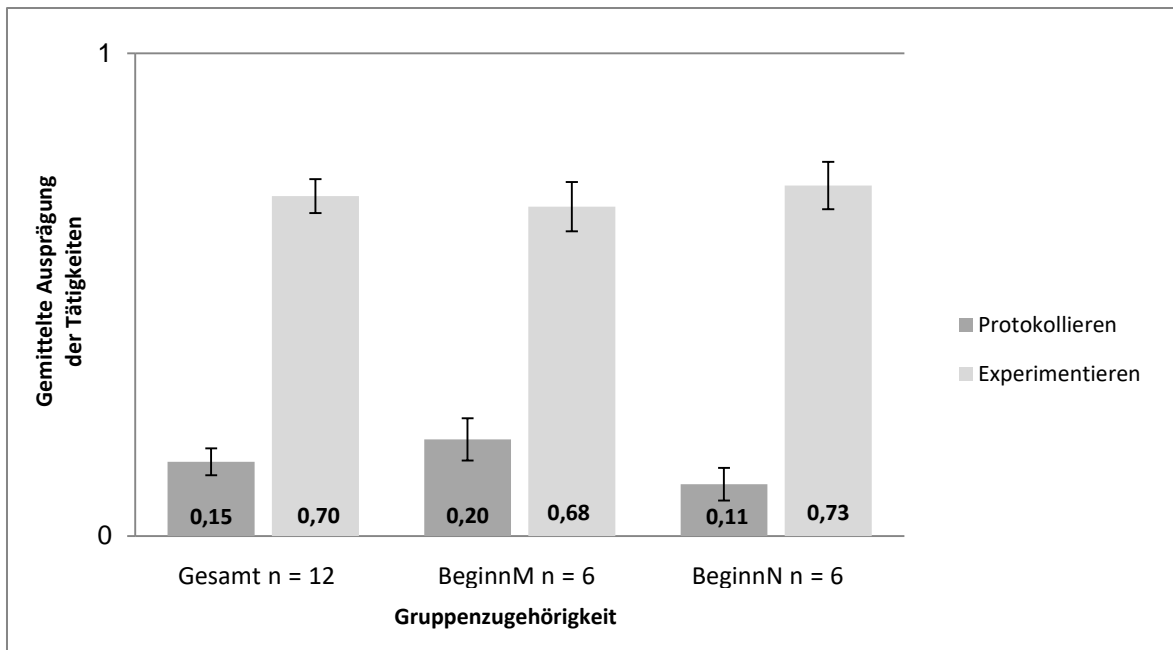


Abbildung 27: Häufigkeit Protokollieren und Experimentieren in Abhängigkeit der Gruppenzugehörigkeit BeginnM und BeginnN mit dem Durchschnitt beider Gruppen als Vergleich. Mittelwerte und Standardfehler der Häufigkeit (0=keine Tätigkeit, 1=Tätigkeit).

Alle nachfolgenden Ergebnisse beziehen sich auf die gesamte Kleingruppe und nicht mehr auf die jeweiligen Protokolle pro Schüler. Die Daten stammen hauptsächlich aus der Videoanalyse, wobei die Wahl der Protokollierhilfe aus den Kategorien der Arbeitsheftanalyse resultierte. Aus Abbildung 28 geht hervor, dass sich die Sozialform beim Protokollieren und Experimentieren unterscheidet. Auf der Skala entsprechen bei den ersten beiden Balken „0“ einer Einzelarbeit und „1“ einer Gruppenarbeit. Ebenso wird erhoben, ob die Protokollierhilfen in der Gruppe gleich sind. Die Null auf der Skala entspricht identischen Protokollierhilfen, die Eins unterschiedlichen Protokollierhilfen. Keine Absprache bei der Wahl der Protokollierhilfe wird durch die Null symbolisiert, eine Eins bedeutet eine Absprache bei der Wahl der Protokollierhilfe. 35% der Probanden protokollierten gemeinsam in der Gruppe. Die Zusammenarbeit nahm beim Experimentieren mit über 40% zu. In der Regel wählte die gesamte Gruppe die gleiche Protokollierhilfe, allerdings erfolgte bei der Auswahl nur in knapp 40% der Fälle eine gruppeninterne Absprache.

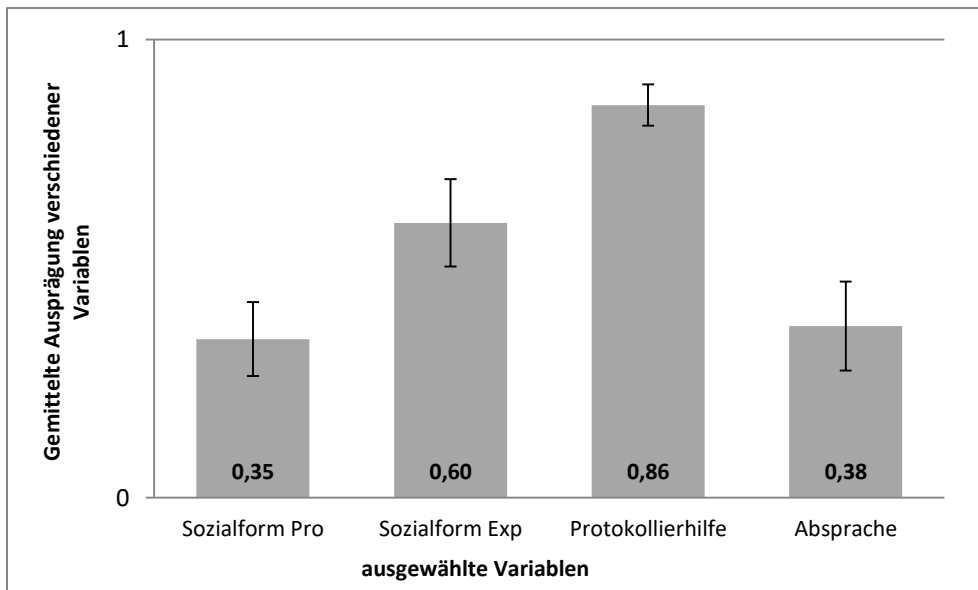


Abbildung 28: Sozialformen beim Protokollieren und Experimentieren, identische Protokollierhilfe sowie Absprache bei der Wahl der Protokollierhilfe dargestellt als Mittelwerte mit Standardfehlern. Variablen von links nach rechts: Sozialform Protokollieren: 0=Einzelarbeit und 1=Gruppenarbeit, Sozialform Experimentieren: 0=Einzelarbeit und 1=Gruppenarbeit, Protokollierhilfe: 0=unterschiedliche Protokollierhilfe und 1=identische Protokollierhilfe, Absprache: 0=keine Absprache und 1=Absprache.

Die Aussagen in den Interviews bestätigen diese Ergebnisse. Die Hälfte der Schüler gibt an, mit ihrem Gruppenpartner beim Protokollieren zusammengearbeitet zu haben. Das Ergebnis liegt in einem vergleichbaren Bereich, wie bei der Videoanalyse.

*T.: „Mit dem...hab' ich die halt auch mit ihm gemacht. Bloß der Deckel hab' ich anders gemalt wie er *kichert*.“*

J.: „Eigentlich hat jeder eher für sich geschrieben. Er hat glaub' ich was anderes.“

Die Sozialform beim Protokollieren hängt mit den heterogenen Voraussetzungen der Schüler zusammen:

J.: „Alleine. Der ist manchmal hinterher oder voraus oder hinterher.“

Trotz der vielen Probleme waren sich alle Befragten des Interviews zumindest in der wichtigen Bedeutung des Protokollierens für die Schule einig:

R.: „Dass man es in Erinnerung behält, dass man's sich immer wieder durchlesen kann und dass man damit auch lernen kann.“

F.: „Da kann man dann...ich war einmal krank, da hat mir jemand das Heft ausgeliehen, da sollte ich des noch nachmachen. Des war eigentlich ganz gut.“

Aber auch in der Forschung ist ein Protokoll notwendig, um Erkenntnisse festzuhalten:

J.: „Ja, weil, vielleicht, falls man da irgendwas Neues entdeckt oder so...dass man des dann aufgeschrieben hat und des nachweisen kann. [...] dann kann ich auch noch seh'n...was ich da gemacht hab. Wenn ich's nicht aufgeschrieben hab', dann war's vielleicht nicht so genau.“

J.: „Deshalb ist ja auch beim Kuchen ein Rezept dabei, sonst ist ja einmal der Kuchen gut und dann kann ihn keiner mehr nachbacken, oder so *kichert*.“

Trotz der anerkannten Bedeutung ist das Protokoll unbeliebt (sechs Aussagen) bis gleichgültig (fünf Aussagen):

J.: „Äh, Spaß macht mir's nicht aber...ein bisschen nervig ist es schon. *grinst* äh...weil ich andauernd Stift rausholen und...manchmal tut mir auch mein Knöchel weh.“

R.: „Pff, nicht schlimm. Langweilig *lacht*.“

V.: „Das ganze Aufschreiben...da muss man halt so viel Schreiben immer!“

Manche Schüler waren sowohl negativ als auch positiv gegenüber Protokollen eingestellt, weshalb mehr Nennungen als interviewte Schüler vorliegen.

J.: Und mitschreiben gehört dazu und ich...es macht auch manchmal Spaß richtig. Also...da zu knobeln, was issen jetzt das genau, warum wird das so...

Nur von knapp 30% der Aussagen konnten positive Äußerung entnommen werden, in denen sich eine gewisse Akzeptanz widerspiegelt.

3.2.5 Diskussion der Ergebnisse

Die Ergebnisse zeigen, dass Schüler unterschiedlich protokollieren und dadurch unterschiedliche Protokolle anfertigen. Die Unterschiede lassen sich durch festgelegte Merkmale am Protokoll (zum Beispiel Wortanzahl), designbedingte Variablen (zum Beispiel Experiment) und interventionsbedingte Faktoren (zum Beispiel Gruppenzugehörigkeit) charakterisieren. Einige dieser Merkmale wurden durch die Analyse der Protokolle und des Prozess des Protokollierens als mögliche Einflussfaktoren auf die Tätigkeit des Protokollierens identifiziert.

Die umfangreiche Wortanzahl bei Experiment 1 und Experiment 6 (Abbildung 20) wird durch die vorgegebene Aufgabenstellung erklärt. Bei Experiment 1 wurden mehrere Stoffkombinationen auf ihre Reaktion getestet, wobei alle Möglichkeiten aufzuschreiben waren. Experiment 6 zielte auf den Vergleich von drei Stoffen ab, die über eine definierte Zeitspanne beobachtet wurden. Dafür waren ebenfalls mehrere Protokolleinträge nötig. Bei Experiment 3 und 4 hingegen schrieben die Probanden im Schnitt 13 Wörter (vgl. Abbildung 20). Beide Experimente behandelten die quantitative Messung von Gasen, die in Experiment 3 identifiziert und in Experiment 4 konzentrationsabhängig verglichen wurden. Die unterschiedlichen Aufgabenstellungen führten zu unterschiedlich ausführlichen Protokollen. Dadurch ist es schwierig die Experimente im Nachhinein untereinander zu vergleichen, da eine äußerliche Gegebenheit auf die Protokolle Einfluss nimmt. Die Wortanzahl der Protokolle unterschiedlicher Experimente zeigen, dass die Aufgabenstellung nicht gleich gestaltet ist. Diese hatten entweder phänomenologischen, messwertbasierten oder kombinatorischen Charakter. Deshalb sind die Experimente und die entstandenen Protokolle sehr unterschiedlich.

Ebenso lässt sich die größere Wortanzahl zusätzlich durch die Wahl der Repräsentationsform begründen. Die Probanden der Kontrollgruppe wählten hauptsächlich Texte als Repräsentationsform, die Experimentalgruppe im Gegenzug eher Tabellen. Texte beinhalten über 40% mehr Wörter und unterscheiden sich so von der Wortanzahl in Tabellen. Die stichpunktartige Aufschriebe oder Kurzsätze der Kontrollgruppe beinhalten in der Regel mehr Wörter als die gewählten Tabellen als Protokollierhilfe der Experimentallgruppe. Tabellen sind grafisch gegliedert, bei denen lediglich die zentralen Ergebnisse als Wort festgehalten werden. In Texten erfolgt die Gliederung durch die Textstruktur, wozu mehrere Wörter notwendig sind. Die Probanden der Untersuchungsgruppe KG mussten ihre Tabellen jedoch selbst anlegen. Diese Arbeit erübrigte sich für die Schüler beim freien Protokollieren.

Die Gestaltung der Arbeitshefte zeigt einen weiteren Unterschied der Protokolle.

Bei den Experimenten 2 und 6 wurde der Protokollierbereich des Arbeitsheftes im Vergleich zu den anderen Experimenten weniger häufig genutzt (vgl. Abbildung 24). In diesen beiden Experimenten waren die Aufgabenstellungen und der Text zur Durchführung auf der Vorderseite zu finden, während sich der Protokollierbereich auf der Rückseite befand. Bei allen anderen Experimenten war der Protokollierbereich auf der gegenüberliegenden Seite verzeichnet, sodass ein Umblättern erspart bleibt. Häufig wurden aufgrund der Arbeitsheftgestaltung unübersichtliche Protokolle auf zu wenig Platz angefertigt. Der Aufbau des Arbeitshefts muss den Schülern übersichtlich und logisch erscheinen. Dabei ist zu berücksichtigen, dass Inhalte chronologisch oder thematisch geordnet werden. Gibt es nicht ausreichend Platz zum Protokollieren oder liegt der Protokollierbereich auf der Rückseite der Aufgabenstellung, werden Protokolle unvollständig oder unübersichtlich verfasst oder nicht angefertigt.

Durch die absoluten Häufigkeiten der Nutzung des Protokollierbereichs zeigt sich indirekt, dass einige Gruppen vor allem Experiment 5 nicht protokolliert haben (vgl. Abbildung 24). Manche Gruppen gerieten gegen Ende der Doppelstunde in Zeitnot, weshalb nicht alle Probanden alle Experimente durchgeführt oder protokolliert haben. Aus diesem Grund zeigt sich jeweils das letzte Experiment einer Doppelstunde als ungeeignet. Die zeitliche Einordnung des Protokolls spiegelt sich in der Tätigkeit des Protokollierens wieder. Auch in den Interviews nannten die Schüler die Zeitnot als Grund für ein nicht fertig gestelltes Protokoll. Bei mangelnder Zeit, besonders gegen Ende einer Lernphase, wurde häufig noch die beliebtere Tätigkeit des Experimentierens abgeschlossen und das Protokollschreiben ausgelassen. Um den Problemen vorzubeugen, muss im Vorfeld der zeitliche Ablauf einer Lernphase gut abgestimmt sein oder währenddessen rechtzeitig eingegriffen werden.

Protokolle haben das Ziel möglichst knapp aber präzise den Weg der naturwissenschaftlichen Erkenntnisgewinnung zu belegen, sodass eine Wiederholung des Inhaltes möglich ist (KRAUS & STEHLIK, 2008). Diese Dokumentation stellt Schüler vor schwierige Herausforderungen (MOLL, 2003; STEINHOFF, 2003; PRECHTL, 2008), was durch die durchgeführten Interviews bestätigt werden konnte. Die Protokollierhilfen sollten die Schüler beim Schreiben unterstützen und sicherstellen, dass Protokolle angefertigt werden. Jedoch gaben die Schüler häufig an mit der Verwendung dieser Hilfsmittel nicht zu Recht gekommen zu sein, da sie die

Konzeption nicht verstanden haben. Die Schüler fühlten sich entgegen der Erwartungen nicht durch die Hilfsmittel unterstützt, sondern hatten eher Probleme mit dem Umgang und der Intention der Vorgaben. Dies gibt Hinweise auf mögliche Fehler bei der Konzeption und den ungeübten Umgang mit den Hilfsmitteln. Genau diese Problematik wurde von den Autoren der Entwicklung von Hilfestellungen beim Protokollieren angemerkt (Kapitel 2.1) und bestätigt sich nun erneut. Unterstützungsmaßnahmen für das Protokollieren sollten vorab auf ihre Nützlichkeit im Allgemeinen untersucht werden, bevor sie als Interventionsmaßnahme eingesetzt werden. Eine Blankoseite hat gegenüber dem Hilfsmittel den Vorteil, dass keine Erwartungen vorgegeben werden, sodass jeder Schüler sein individuelles Protokoll in Abhängigkeit seiner Fähigkeit anfertigen kann. Zwar besteht die Gefahr, dass die Schüler sich so auf Nebensächlichkeiten konzentrieren, allerdings ist die Eigenleistung beim Protokollieren höher. Eine Blankoseite fordert mehr Kreativität und die Schüler müssen über das eigene Handeln mehr nachdenken. Auch die Absprache bei der Wahl der Protokollierhilfe erfolgt häufig ohne Einbindung der Gruppe. Oft wählte ein einzelner Schüler die Protokollierhilfe für die gesamte Gruppe, weshalb sich die Variablen Absprache und Protokollierhilfe zu über 55% unterschieden. Weitere Probleme, wie das hohe Anforderungsniveau, Unterbrechungen im Lernfluss oder mangelnde Motivation störten Schüler neben der Arbeitsheftgestaltung und dem Zeitfaktor beim Protokollieren (vgl. Interviews Seite 57).

Die Inhalte dieser Teilstudie wurden in Kleingruppen erarbeitet. Gruppen entwickelten automatisch eine gewisse Gruppendynamik, da sich eine Gruppe aus einer heterogenen Schülergemeinschaft zusammensetzte. Innerhalb einer Gruppe kristallisieren sich häufig einzelne Rollen heraus, beispielsweise der Anführer, der Vermittler oder der Initiator. Die Gruppendynamik schlägt sich in der Tätigkeit des Protokollierens nieder. Die Unterschiede in der Sozialform beim Protokollieren und Experimentieren (vgl. Abbildung 28) sind mit den zur Verfügung gestellten Materialien erklärbar. Während jeder Schüler ein eigenes Arbeitsheft ausfüllte, gab es nur einen Versuchsaufbau für eine Kleingruppe. Damit das Experiment gelang, mussten die Probanden zusammenarbeiten. Ab und zu war eine Person abgelenkt und nicht am Experiment beteiligt. Die Protokolle wurden häufig in Einzelarbeit verfasst, eine Gruppenarbeit kam nur zustande, wenn die Schüler Ergebnisse verglichen oder Hilfe in der Kleingruppe suchten. Die unterschiedlichen Lerntempi der Schüler verhinderten eine optimale Zusammenarbeit, wodurch die Tätigkeit des Protokollierens häufiger in Einzelarbeit voll-

bracht wurde als die Tätigkeit des Experimentierens. Die Interviews zeigen, dass die Gruppendynamik von der Heterogenität innerhalb einer Kleingruppe beeinflusst wird.

Die Unterschiede in der Wortanzahl und Verwendung des Protokollbereichs, die Probleme bei dem Einsatz der Hilfsmittel und die Unterschiede durch das gemeinsame Arbeiten beim Protokollieren zeigen, dass die Gestaltung der Lerngegebenheiten Unterschiede in den Protokollen hervorruft.

Durch die Interviews konnte gezeigt werden, dass die Schüler Vorerfahrung im Protokollieren haben, die Kenntnisse über ein Protokoll jedoch sehr unterschiedlich sind. In den Untersuchungsgruppen BeginnN und BeginnM wurden signifikante Unterschiede in der Wortanzahl der Protokolle aufgezeigt. Während die Schüler „BeginnN“ in der Nawi-Werkstatt zum ersten Mal in Kontakt mit außerschulischem Protokollieren gekommen waren, konnten die Schüler „BeginnM“ im Mathematik-Labor schon Vorerfahrungen sammeln. Die Kenntnisse trugen die Probanden in die anschließende Intervention in der Nawi-Werkstatt mit und schrieben deshalb ausführlichere Protokolle. Dies gibt Auskunft über Auswirkung der Vorkenntnisse auf das Protokollieren. Die Probanden, die zuerst das Mathematik-Labor besuchen, protokollierten im Schnitt ausführlicher und häufiger, als die Schüler, die ohne Vorkenntnisse mit dem außerschulischen Lernort in der Nawi-Werkstatt begannen. Diese Unterschiede in den Protokollen von BeginnM und BeginnN sind unabhängig vom Einsatz der Protokollierhilfe oder von der Vorerfahrung in der Schule. Es gibt keinen weiteren Zusammenhang der Variable, beispielsweise mit der Nawi-Schulnote. Die zwei Experimentalgruppen sind also vergleichbar. Alles deutet darauf hin, dass die Schüler am ersten außerschulischen Lernort gelernt haben zu protokollieren und diese Vorkenntnisse für den zweiten außerschulischen Lernort mitbrachten. Allerdings protokolliert jedes Fachgebiet anders. Zwar ähneln sich die Protokolle der naturwissenschaftlichen Fächer Biologie, Chemie und Physik, jedoch gibt es Unterschiede zum Fach Mathematik. Trotz dieser Unterschiede zwischen Mathematik und dem Fach Naturwissenschaften erlernten die Probanden durch diese Vorkenntnisse bereits die Fertigkeit Protokolle anzufertigen. Es könnte dadurch abgeleitet werden, dass diese Fertigkeit fachunabhängig ist. Jedoch deckt diese empirische Studie nur einen geringen Teil dieser Aussage ab und für eine solche Erkenntnis wären weitere Untersuchungen notwendig. Weitere signifikante Unterschiede in der Wortanzahl zeigt die Nawi-Schulnote auf: Sehr gute Schüler verwenden wenige Worte um ein Protokoll zu verfassen, können sich also gezielter ausdrücken. Schüler mit ausreichenden Nawi-Kenntnissen schreiben vermutlich auch unnö-

tige Details auf beziehungsweise fehlt ihnen die Fähigkeit, ihre Aussagen zu konkretisieren. Nicht nur die Vorkenntnisse im Protokollieren sondern auch die Kenntnisse und Leistungen des Faches zeigen somit einen Einfluss auf die Protokolle.

Alle interviewten Schüler waren sich der Relevanz des Protokollierens bewusst, sowohl für die Wissenschaft als auch für die Schule. Es stellt sich die Frage, warum das Protokollieren eine derart unbeliebte Tätigkeit ist, obwohl die Schüler um dessen Bedeutung nur zu gut Bescheid wissen. In den Äußerungen der Schüler zeigt sich, dass die Genauigkeit der naturwissenschaftlichen Arbeitsweise und die ausführliche Dokumentation bei Schülern Schwierigkeiten verursacht, die sich negativ auf die Einstellung zum Protokollieren auswirkt. Die Schüler bringen Protokolle in Verbindung mit der notwendigen Dokumentation von Ergebnissen, damit diese replizierbar werden. Umso gravierender ist die Tatsache, dass die Einstellung zum Protokollieren negativ belegt ist. Die wohl größte Unbeliebtheit geht mit der mühevollen Schreibtätigkeit einher. Es ist davon auszugehen, dass eine unbeliebte Tätigkeit nicht mit einer hohen Intensität durchgeführt wird. Ein Schüler der unmotiviert protokolliert, tut dies unbewusst wahrscheinlich schlechter, als ein Schüler, der mit einer hohen Motivation protokolliert. Für eine genauere Aussage, müsste jedoch die Schreibmotivation erfasst werden und mit der Protokollqualität verglichen werden. KRAUS & STEHLIK (2008) konnten ebenso aufzeigen, dass der Protokollbegriff in seiner Notwendigkeit über die Jahrgänge 7-12 den Schülern bewusst ist und trotzdem bei den Befragten eine negative Assoziation hervorruft.

Die Beantwortung der Forschungsfragen zeigte, dass sowohl im Produkt Protokoll als auch im Prozess des Protokollierens Unterschiede zu tragen kommen. Der Grund dafür liegt in verschiedenen Faktoren die sich auf das Protokoll und die Tätigkeit des Protokollierens auswirken. Es lassen sich fünf übergeordnete Kategorien zusammenfassen, welche Einfluss auf die Protokollemerkmale und das Protokollieren nehmen.

- 1) Experiment: Der Bereich „Experiment“ bezieht das Anforderungsniveau der Aufgabenstellung, die inhaltliche Einordnung, sowie den phänomenologischen, messwertbasierten oder kombinatorischen Charakter des Versuchs mit ein. Die Art des Experimentes übt einen Einfluss auf die geschriebene „Wortanzahl“ aus.
- 2) Lerngegebenheiten: Der Kreis „Lerngegebenheiten“ umschließt die Auswirkungen der Lernumgebung, den Zeitfaktor und alle zur Verfügung gestellten Materialien, beispielsweise das Arbeitsheft oder die Art der Protokollierhilfe. Die Lerngegebenheiten stehen in Beziehung mit den Variablen „Wortanzahl“ und „Protokollierbereich“. Im Gegensatz zum Experiment sind die Lerngegebenheiten unabhängig von dem zu vermittelnden Inhalt.
- 3) Vorkenntnisse: Das Feld „Vorkenntnisse“ umfasst die Ausprägung der fachunabhängigen Fähigkeit zu protokollieren durch die bisherige Leistung eines Schülers, beispielsweise der Note im Fach Naturwissenschaften und Vorerfahrungen in der Schule oder an einem außerschulischen Lernort. Die Vorkenntnisse beeinflusst die Variable „Wortanzahl“.
- 4) Gruppendynamik: Der Ring „Gruppendynamik“ beinhaltet Einflüsse auf das Protokollieren, wie die Rollenverteilung, Heterogenität oder Interaktion innerhalb einer Kleingruppe. Der Einfluss der Gruppendynamik zeigt sich zum Beispiel in den Variablen „Sozialform beim Protokollieren“, „Sozialform beim Experimentieren“ und „Absprache“.
- 5) Einstellung: Die Kategorie „Einstellung“ involviert personenabhängige Neigungen beim Protokollieren, zum Beispiel die eigene Motivation oder die Vorliebe für eine bestimmte Art des Protokollierens. Ein Zusammenhang zwischen Einstellung und Protokoll wurde nicht direkt überprüft.

3.2.6 Methodenkritik und Änderungsvorschläge

Probleme und Auffälligkeiten dieser Studie werden in diesem Kapitel erläutert und Verbesserungsvorschläge angebracht. Auf Grundlage dieser Teilstudie wurde eine weitere Studie initiiert. Teilstudie 2 „Schreiben von Protokollen“ brachte Erkenntnisse über verschiedene Einflussfaktoren, welche sich auf das Protokoll als Produkt und auf die Tätigkeit des Protokollierens auswirken und Unterschiede hervorrufen. Diese Faktoren sind bei der Konzeption weiterer Studien zu beachten, wenn ein vergleichbares Protokoll oder Tätigkeit des Protokollierens im Fokus stehen soll. Die Erkenntnisse und Änderungsvorschläge dieses Kapitels flossen bei der Konzeption von Teilstudie 3 „Bedeutung von Protokollen“ ein.

Forschungsfrage und Design

Diese Studie hatte einen ausgeprägten explorativen Charakter. Eine solche Herangehensweise war jedoch notwendig, um sich diesem Forschungsfeld zu nähern. Dadurch war es schwierig Forschungsfragen zu formulieren und zusätzlich präzise zu beantworten. Teilweise wurde das Design aufgrund von organisatorischen Rahmenbedingungen der Forschungsinitiative vor der möglichen Fragestellung generiert. Deshalb wurden im Anschluss nur allgemeine Forschungsfragen ohne Hypothesen formuliert. Das Problem lässt sich am Design der Studie festmachen: Für die Beantwortung der Forschungsfragen waren weder ein Gruppenvergleich noch drei Messzeitpunkte nötig. Das Design bedingte verschiedene Vergleiche, die für den Einfluss auf das Protokollieren nicht notwendig sind. Dieses Problem zeigt sich ebenso in der Auswahl der Erhebungsinstrumente und der Auswertungsmethoden. Die Auswahl ist sehr unspezifisch für die Beantwortung einer Forschungsfrage. Eine genauere Ausarbeitung der Methoden der Datenerhebung und Datenauswertung wäre durch eine Konkretisierung mit Hypothesen möglich gewesen. Die Kooperation mit verschiedenen Fächern hat dieses Vorgehen notwendig gemacht. Nur so konnte interdisziplinär dieses Thema bearbeitet werden.

Die Untersuchung hat ebenso gezeigt, dass Schüler protokollieren, wenn sie die Möglichkeit zum Schreiben haben. Es wurde im Gegenteil zuvor angenommen, dass die Schüler in der Kontrollgruppe ohne Protokollierhilfen keine Protokolle anfertigen werden würden. Auch wenn bei Schülern die Tätigkeit des Protokollierens unbeliebt ist, nutzen sie aufgrund der gegebenen Strukturen (leere Seite und Piktogramm) die Möglichkeiten zum Schreiben. Streng genommen wird durch dieses Design nur die Qualität der Hilfsmittel verglichen. Soll

der direkte Vergleich zwischen dem Anfertigen eines Protokolls und dem fehlenden Protokollieren angestellt werden, muss das Design und die Intervention angepasst werden.

Erst nach Beendigung der Studie fiel den beteiligten Personen auf, dass die zuständigen Lehrer die Gruppeneinteilung verändert hatten. Deshalb war die Gruppeneinteilung nicht durchweg konstant. Manche Klassen waren nur in sieben anstatt acht Gruppen aufgeteilt. Dadurch war die Verteilung von Experimental- und Kontrollgruppe nicht so einheitlich wie angenommen.

Methoden der Datenerhebung

Der Fachtest wurde zwar nicht ausgewertet, soll jedoch an dieser Stelle trotzdem diskutiert werden, um die Vorgehensweise für eine weitere Studie zu verbessern. Die Aussagekraft des Leistungstests ist bei 13 Items mit unterschiedlichen Antwortformaten (Multiple Choice, offene Aufgaben mit Textfeldern, richtige Antworten mit Stichen verbinden) als gering einzustufen. Teilweise wurde ein Experiment mit nur einem Item überprüft. Der Entwicklung von ausreichend Fach-Items wurde zu wenig Gewicht zugeordnet. Es sind mehr Fragen insgesamt und pro Experiment notwendig, um eine genauere Aussage über die Lernleistung während der Intervention treffen zu können. Aussagen über die Lernleistung im Gruppenvergleich wurden deshalb nicht aufgestellt, obwohl es im Voraus angedacht war. Nachgehende Analysen stützen dieses Problem: Zum Beispiel wurde durch mehrere Varianzanalysen überprüft, ob das Ergebnis einer Aufgabe durch verschiedene Variablen beeinflusst werden. Ein richtiges Ergebnis im Post-Test der Aufgabe 10, kann mit einer Fehlerwahrscheinlichkeit von $p=0.012^*$ durch die Note im Fach Naturwissenschaften und $p=0.004^{**}$ durch das Ergebnis im Pre-Test erklärt werden. Dieses Ergebnis zeigt, dass der Leistungstest bereits durch eine gute Note oder ein hohes Vorwissen beantwortet werden kann und nicht durch eine Leistung während der Intervention.

Konzeption der Lerneinheit

Die Lerneinheit war nach Aussage der beteiligten Lehrkräfte in sich thematisch und methodisch sinnvoll aufgebaut und durch das eigenständige Arbeiten motivierend für die Schüler. Die komplexe Struktur mit unterschiedlichen Hilfestellungen und die ausführlichen Experimente erwiesen sich jedoch für Teilnehmer der sechsten Klassenstufe zu schwierig. Während der Einheit suchten die Schüler oft eine Hilfestellung durch die Betreuer. Das erwünschte

Ergebnis der Experimente wurde in manchen Fällen durch ein falsches Vorgehen oder unkonzentriertes Lesen der komplexen Anleitungen nicht erzielt. Die ungleiche Gestaltung der Experimente macht diese untereinander und dadurch auch die Ergebnisse nicht vergleichbar. Eine klar strukturierte Stundengestaltung sowie eine gleiche Anzahl an Experimenten würde die Lerneinheit transparenter machen. Um alle Experimente vergleichen zu können, sollten diese einer vergleichbaren Struktur und Aufgabenstellung folgen. Auf die Hilfestellungen beim Experimentieren durch Fotostories kann verzichtet werden, jedoch müssen die Experimente gelingen und eindeutige Ergebnisse aufweisen.

Methoden der Datenauswertung

Es existieren wenige Literaturhinweise zur Bewertung von Protokollen (zum Beispiel EMDEN & SUMFLETH, 2012). Deshalb wurden in diesem Projekt durch die qualitative Inhaltsanalyse erste Anhaltspunkte ermittelt, wie sich Protokolle auswerten lassen. Manche Kategorien erwiesen sich im Nachhinein jedoch nicht als hilfreich zur Beantwortung der Fragestellung und wurden deshalb nicht dargestellt. Das Problem bei der Auswertung der Protokolle war deren fehlende Vergleichbarkeit. Um weitere Protokolle über eine gleiche Auswertungsmethode zu bewerten, sollte eine identische Aufgabenstellung benutzt werden. Für eine genauere Analyse des Lernerfolgs in Bezug auf das Protokoll, wäre es sinnvoll, wenn das Wissen über die einzelnen Experimente sich auf die jeweils passenden Protokolle beziehen würde.

Um die Aussagekraft der Ergebnisse zu erhöhen, muss insgesamt statistisch genauer gearbeitet werden. Dafür sollten zunächst die Voraussetzungen des Datenmaterials geprüft werden (zum Beispiel Normalverteilung und Varianzhomogenität der Daten). Anschließend ist eine systematische Suche nach Zusammenhängen und Unterschieden der Variablen durchzuführen. Dabei ist auch die genaue Verwendung der passenden Testverfahren zu achten. Dafür müssen die Variablen genau definiert und erfasst werden. Ebenso sind vorab genauere Überprüfungen des Fachtests durch Itemanalysen notwendig, um dessen Messgenauigkeit zu testen und die Items möglicherweise zu überarbeiten. Eine Itemanalyse war in dieser Studie aufgrund der geringen Itemzahl pro Experiment irrelevant.

3.3 Empirische Studie zur Bedeutung von Protokollen – Lernwirksamkeit und Qualität

Teilstudie 2 zeigt, wie Schüler protokollieren und welche Einflussfaktoren auf diesen Prozess einwirken können. Durch die theoretischen Überlegungen wurde die Bedeutung des Protokollierens seitens der Bildungsstandards begründet. Jedoch erfolgte seitens der Literatur noch keine empirische Legitimation für das Schreiben von Protokollen. Zwar wird der Prozess des Schreibens in den Naturwissenschaften generell als lernförderlich erachtet und wird auch durch mehrere Studien belegt, jedoch gab es keinen Übertrag dieser Theorie bezüglich des Protokolls. Das Bildungswesen erhofft sich grundsätzlich von einer Tätigkeit, wie dem Protokollieren, einen Ertrag für die Lernwirksamkeit.

Teilstudie 3 basiert im Wesentlichen auf der Methodenkritik von Teilstudie 2. Ausgehend von der theoretischen Grundlage zum naturwissenschaftlichen Erkenntnisprozess, der Bedeutung des Schreibens und der Ergebnisse der vorangegangenen Studie stellen sich zwei Forschungsfragen, um die Bedeutung des Protokollierens und dessen Wirkung genauer zu erfassen.

➤ Forschungsfrage 2.1

Welche Bedeutung hat das Protokollieren für den Lernerfolg?

Das Ziel eines Lehrenden ist es, eine hohe Lernwirksamkeit der Schüler durch den Lerngegenstand zu erreichen. Das Protokoll ist eine gängige Methode im Unterricht für die Sicherung des Experimentierprozesses. Es wird davon ausgegangen, dass das Anfertigen eines Protokolls während des Experimentierens den Lernerfolg erhöht. Jedoch wurde diese Hypothese noch nicht untersucht.

(H1) Das Anfertigen eines Protokolls während des Experimentierens erhöht den Lernerfolg der Schüler im Vergleich zu Schülern, die kein Protokoll anfertigen.

➤ Forschungsfrage 2.2

Welche Bedeutung hat die Qualität des Protokolls für den Lernerfolg?

Angenommen Hypothese 1 kann verifiziert werden, stellt sich die Frage welche Bedeutung die Qualität eines Protokolls auf den Erwerb der Lernwirksamkeit darstellt. Schüler werden dazu gedrängt ein vollständiges und strukturiertes Protokoll anzufertigen. Jedoch wurde bis-

her noch nicht evidenzbasiert nachgewiesen, dass nur ein „gutes Protokoll“ zum Lernerfolg führt. Ist möglicherweise bereits der reine Schreibvorgang entscheidend für den Lernerfolg? Um diese Fragen beantworten zu können, war ein Messinstrument erforderlich, mit der Protokolle auf ihre Qualität überprüft werden können. Ein solches Instrument war in der Literatur nicht zu finden. Zunächst musste geklärt werden, welche Kriterien für die Auswertung von Protokollen herangezogen werden können. Die Zusammenhänge zwischen der Qualität der Protokolle und dem daraus resultierenden Lernerfolg eines Experimentes soll empirisch geprüft werden. Die Hypothesen zu dieser Forschungsfrage lauten:

(H2) Die Lernwirksamkeit, basierend auf einer experimentellen Tätigkeit, korreliert mit der Qualität des Protokolls, welches dazu angefertigt wurde.

(H3) Schüler, die ein gutes Protokoll anfertigen, weisen einen höheren Lernerfolg auf als Schüler, die ein schlechtes Protokoll erstellen.

Design und Stichprobe

Zur Beantwortung der Hypothesen wurde eine empirische Studie im Schülerlabor Nawi-Werkstatt durchgeführt, welche von Februar bis August 2014 stattfand (vgl. Abbildung 29). Zunächst wurden die eingesetzten Unterrichts- und Testmaterialien pilotiert und anschließend überarbeitet. Als Untersuchungsdesign wurde die Methode der Vergleichsstudie mit zwei Interventionsgruppen (Experimentalgruppe und Kontrollgruppe) und drei Messzeitpunkten (Pre, Post und Follow-Up) gewählt. Jede Gruppe besuchte das Schülerlabor Nawi Werkstatt über drei Wochen hinweg für je einen Termin pro Woche. Der Pre-Test wurde in der Woche vor der Intervention und der Post-Test in der Woche nach der Intervention durchgeführt. Da nicht alle Klassen gleichzeitig die dreiwöchige Intervention durchführten, erstreckte sich der zeitliche Rahmen der Durchführung über zwei Monate. Jeweils drei Monate nach dem Post-Test fand der Follow-Up-Test statt. Die Lernwirksamkeit gilt als zentrale abhängige Untersuchungsvariable für die Untersuchung beider Forschungsfragen. Die Zugehörigkeit zu der Interventionsgruppe definiert die unabhängige Variable von Forschungsfrage 2.1 und die Qualität der Protokolle stellt die unabhängige Variable von Forschungsfrage 2.2 dar.

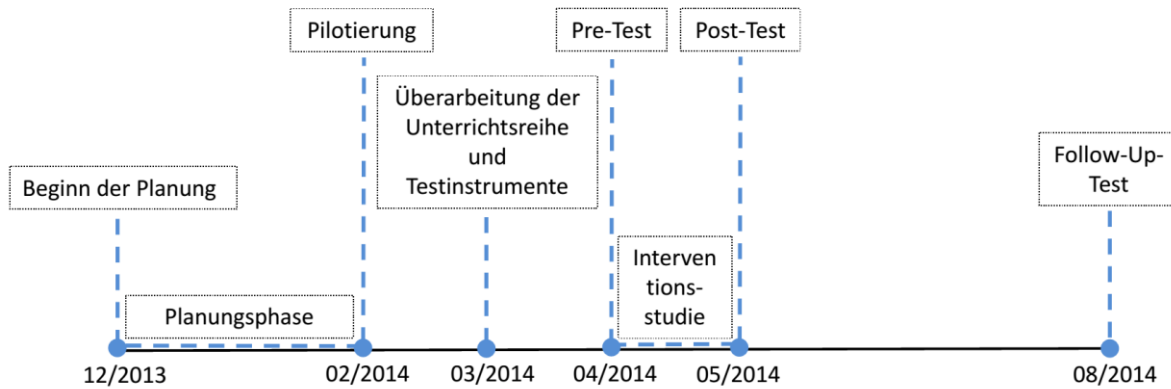


Abbildung 29: Zeitlicher Verlauf der empirischen Studie.

Die Stichprobe umfasste vier Parallelklassen der sechsten Jahrgangsstufe des Max-Slevogt-Gymnasiums in Landau. Auch hier entschieden wir uns wieder bewusst für die sechste Jahrgangsstufe, in der das Protokollieren noch nicht eingeführt wurde. Dadurch wurde ein unterschiedlicher Kenntnisstand vermieden. Von den insgesamt 115 beteiligten Schülern waren 41% weiblich und 59% männlich. Es ergab sich eine Stichprobe aus $n=55$ für die Kontrollgruppe und $n=60$ für die Experimentalgruppe. Die Schüler waren im Durchschnitt 11,2 Jahre alt. 23% der Schüler weisen einen Migrationshintergrund auf.

Um die Lernervariable (Merkmale und Voraussetzungen von Schülern) möglichst zu kontrollieren, fand die Einteilung der Kontroll- und Experimentalgruppe durch einen Aufbruch des Klassenverbandes statt (vgl. Abbildung 30). Die Gruppeneinteilung wurde aus organisatorischen Gründen von den Lehrern vorgenommen und war deshalb quasi-randomisiert. Die beiden Interventionsgruppen arbeiteten jeweils in getrennten Räumen.

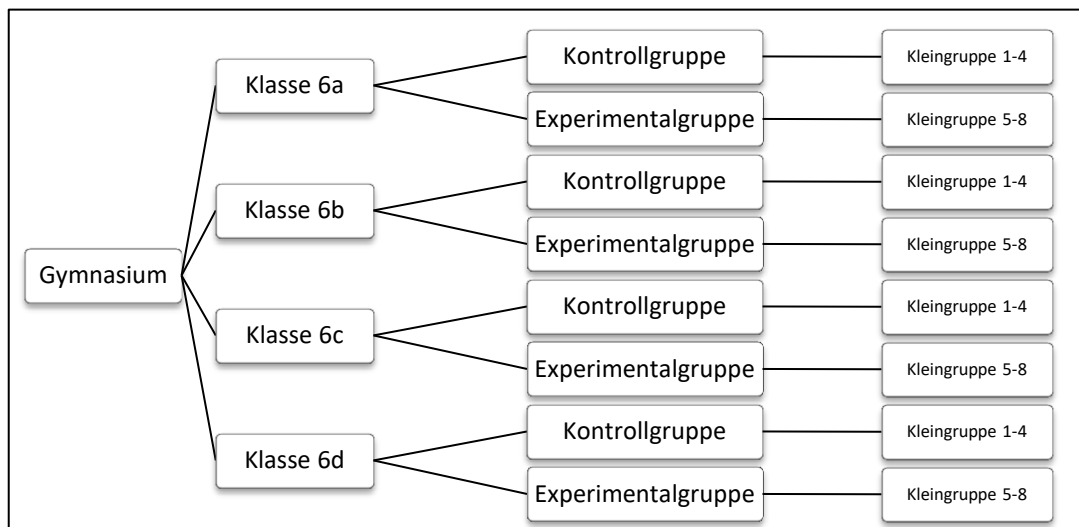


Abbildung 30: Einteilung der Vergleichsgruppen.

Innerhalb einer Interventionsgruppe arbeiteten die Schüler in einer Kleingruppen von drei bis vier Personen. Daraus ergaben sich acht Kleingruppen pro Klasse, wovon je vier der Kontrollgruppe und vier der Experimentalgruppe zugeteilt wurden (vgl. Abbildung 30). Somit ergaben sich insgesamt 16 Kleingruppen in der Kontrollgruppe und 16 Kleingruppen in der Experimentalgruppe.

Interventionsmaßnahmen

Für einen fairen Vergleich der Lerneffekte der Interventionsmaßnahme war es wichtig, die Einflussfaktoren ausreichend zu kontrollieren und möglichst gleiche Bedingungen in den Gruppen zu generieren (THEYSEN, 2014). Die Inhalte und Lerngegebenheit wurden durchweg gleichwertig gestaltet: Beide Gruppen erhielten die identische Aufgabenstellungen, Ziele und Experimentiermaterialien. Eine naturwissenschaftliche Fragestellung konnte mit Hilfe der vorhandenen Materialien und der Versuchsbeschreibung bearbeitet werden (siehe Konzeption der Lerneinheit Kapitel 3.3.3). Die Strukturierung der Bearbeitungszeit war bei beiden Gruppen an die in Kapitel 3.1 definierten angepasste Struktur des Erkenntniswegs angelehnt. In jeder der drei Doppelstunden wurden zwei Experimente à 40 Minuten bearbeitet (Abbildung 31). Die Interventionsmaßnahme wurde in den Ablauf eines jeden Experimentes integriert. Die Interventionsmaßnahme für die Vergleichsgruppen wurde durch zwei Teilaspekte umgesetzt. Die Intervention der Experimentalgruppe bestand in der Aufforderung ein Protokoll über das durchgeführte Experiment anzufertigen. Dazu wurde eine Protokollphase (fünf Minuten pro Experiment) in den Ablauf integriert und die Gestaltung des Arbeitsheftes im Vergleich zur Kontrollgruppe geändert.

Kontrollgruppe

Fragestellung	Experiment	Schlussfolgerung
5 Minuten	30 Minuten	5 Minuten

Experimentalgruppe

Fragestellung	Experiment	Daten aufschreiben & auswerten	Schlussfolgerung
5 Minuten	25 Minuten	5 Minuten	5 Minuten



In Form eines Protokolls

Abbildung 31: Zeitlicher Verlauf pro Experiment und Interventionsmaßnahme.

Die Probanden der Experimentalgruppe erhielten im Arbeitsheft die Aufforderung, die Ergebnisse des Experimentes zu verschriftlichen. Im Vergleich dazu wurde in der Kontrollgruppe das Schreiben von Protokollen verhindert. Die Probanden erhielten keinen Stift und ein Arbeitsheft, das keine Protokolliermöglichkeit aufwies (vgl. Abbildung 32). Um die häufig kritisierten Randbedingungen Neuigkeitseffekt und Lehrervariable zu kontrollieren, besuchten die Schulklassen mehrfach das Schülerlabor und arbeiteten mit Hilfe des Arbeitsheftes weitgehend selbstständig. Um die Bearbeitungszeit in beiden Gruppen gleich zu halten, wurde in der Kontrollgruppe die Protokollphase durch eine längere Experimentierzeit kompensiert.

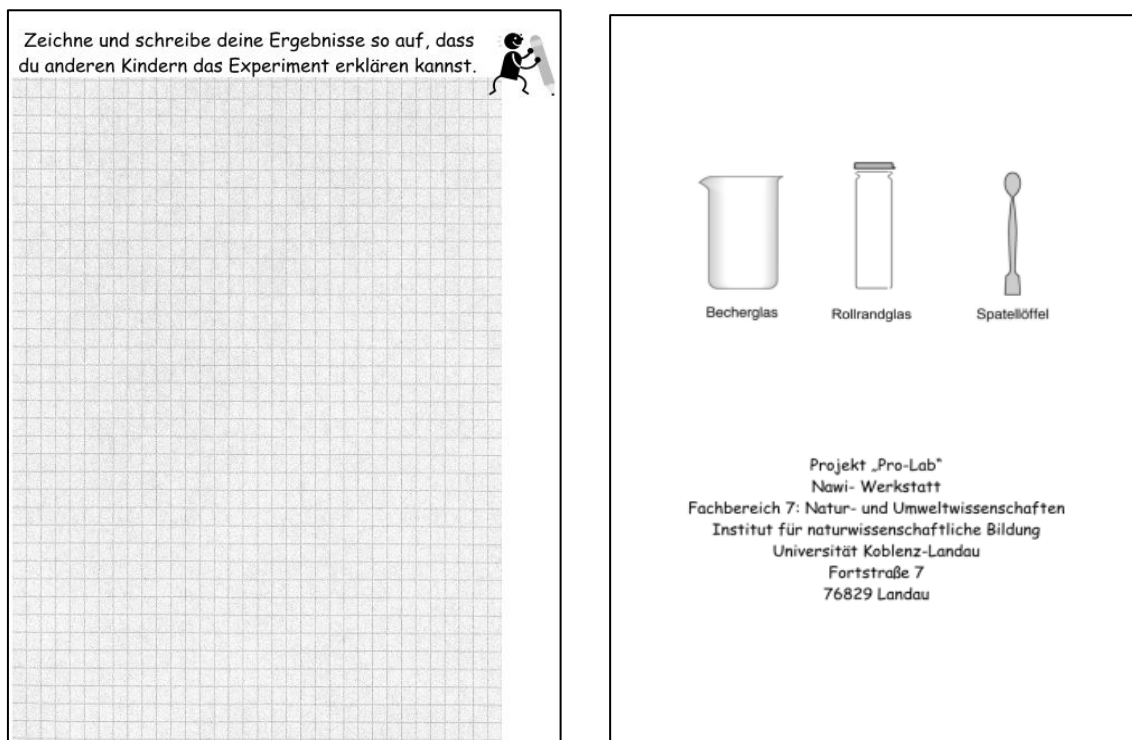


Abbildung 32: Gestaltung der Arbeitshefte: links die Seite aus dem Heft der Experimentalgruppe, rechts die Seite aus dem Heft der Kontrollgruppe.

3.3.1 Pilotierung

Die Lerneinheit und der Wissenstest wurden vor Beginn der Vergleichsstudie mit einer sechsten Klasse (N=24) des Goethe-Gymnasiums in Neustadt an der Weinstraße pilotiert. Diese Klasse erprobte im Schülerlabor die ausgearbeitete Lerneinheit und Intervention, sowie Pre- und Post-Test.

Bei der Überprüfung der Lerneinheit standen die Verständlichkeit der Arbeitshefte und Fragestellungen sowie die zeitliche Organisation der Experimente im Fokus. Durch die Beobachtung der Schüler während des Experimentierens und die anschließende Besprechung der Einheit mit den betreuenden Lehrkräften wurden Schwachstellen der Untersuchung identifiziert und im Anschluss überarbeitet. Probleme im Stundenablauf und bei den Experimenten wurden aufgrund der Pilotierung bis zur Hauptstudie verbessert.

Pilotierung des Leistungstests

Bei der Entwicklung von Leistungstests galt es zu beachten, dass die Aufgaben eindeutig formuliert und sich sowohl inhaltlich als auch sprachlich dem Anforderungsniveau der Probandengruppe entsprechen (RIESE & REINHOLD, 2014). Sprachlich sollten unnötige Fremdwörter, doppelte Verneinungen und mehrdeutige Begriffe vermieden werden (ebd.). Der Leistungstest wurde auf sprachliche Verständlichkeit mit der Klasse besprochen und Missverständnisse für die Überarbeitung notiert. Die Bearbeitungszeit der Tests (45 Minuten pro Testzeitpunkt) war ausreichend damit jeder Proband das Testheft bearbeiten konnte. Anschließend wurde eine Itemanalyse des Post-Tests der Pilotierungsklasse durchgeführt. Zu schwierige oder zu leichte Items wurden mit Hilfe der Häufigkeitsverteilung der Antwortmöglichkeiten identifiziert und anschließend verändert oder aus dem Test rausgenommen, um deutliche Boden- und Deckeneffekte zu vermeiden (RIESE & REINHOLD, 2014). Die Veränderung der Items erfolgte durch das Umformulieren von Aufgaben oder Veränderung der Attraktoren oder Distraktoren. Zehn Items von 31 Items aus dem Fachtest wurden gestrichen und bei fünf erfolgte eine Überarbeitung der Distraktoren. Aufgrund der vielen Ausfällen von Items wurden anschließend sieben neue Items entwickelt. Die Tabelle 9, Tabelle 10, Tabelle 11 stellen die Vorgehensweise der Bearbeitung beispielhaft dar.

Tabelle 9 zeigt ein Beispiel für eine Häufigkeitsverteilung der Antwortmöglichkeiten, welche zu der Entscheidung führte, das Item weiter im Test zu verwenden. Alle Antwortmöglichkeiten wurden genutzt. Die richtige Antwort (Nr. 4) wurde mit über 50% gewählt, aber auch

die Distraktoren wurden alle gewählt. Die mittlere relative Lösungshäufigkeit sollte nach RIESE & REINHOLD (2014) zwischen 20% und 80% liegen.

Tabelle 9: Häufigkeitsverteilung der Nennung der Antwortmöglichkeit eines Items für Experiment 2.

Antwortmöglichkeit	Häufigkeit der Nennung	Prozentuale Häufigkeit der Nennung
1	1	5 %
2	6	32 %
3	2	11 %
4	10	53 %

Tabelle 10 zeigt ein Beispiel für eine Häufigkeitsverteilung der Antwortmöglichkeiten, welche zur Veränderung der Distraktoren führte. Nicht alle Antwortmöglichkeiten wurden genutzt, was darauf hinweist, dass diese Distraktoren zu trival formuliert waren. Die richtige Antwort (Nr. 1) wurde von fast allen Schülern erkannt, was auf ein zu einfaches Item schließen lässt.

Tabelle 10: Häufigkeitsverteilung der Nennung der Antwortmöglichkeit eines Items für Experiment 6.

Antwortmöglichkeit	Häufigkeit der Nennung	Prozentuale Häufigkeit der Nennung
1	16	84 %
2	2	11 %
3	1	5 %
4	0	0 %

Tabelle 11 zeigt ein Beispiel für die Häufigkeitsverteilung der Nennung der Antwortmöglichkeiten, welche zum Entfernen des Items aus dem Test geführt hat. Der Attraktor (Nr.3) wurde von nahezu allen Schülern genannt. Nur ein Distraktor wurde gewählt, zwei weitere Distraktoren wurden nicht gewählt. Das Item mit den Antwortmöglichkeiten ist zu leicht und zeigt wenig Ansatz für Überarbeitungsmöglichkeiten.

Tabelle 11: Häufigkeitsverteilung der Nennung der Antwortmöglichkeit eines Items für Experiment 3.

Antwortmöglichkeit	Häufigkeit der Nennung	Prozentuale Häufigkeit der Nennung
1	0	0 %
2	1	5 %
3	18	95 %
4	0	0 %

Die Pilotierung sollte laut Literatur mit mindestens 40 Personen stattfinden, um klassische psychometrische Kennwerte wie Trennschärfe und Reliabilität berechnen zu können (RIESE & REINHOLD, 2014). Da dies nicht gegeben war, wurde für die Pilotierung keine Reliabilitäten oder Trennschärfen bestimmt. Eine ausführliche Überprüfung der Items auf diese Kriterien fand deshalb erst nach der Pilotierung statt.

3.3.2 Methoden der Datenerhebung

Die Daten wurden zu drei Hauptmesszeitpunkten und begleitend zur Intervention erhoben. Den Schülern wurde während der Testphasen und der Intervention eine individuelle Identifikationsnummer zugeteilt. Die sechsstellige Nummer bestand aus vier schulinternen Ziffern plus zwei Ziffern des Geburtstages. Dies ermöglichte eine anonyme individuelle Rückverfolgung der Lernleistung (TIEMANN & KÖRBS, 2014). Zusätzlich wurden im Pre-Test (Testheft siehe Anhang V) allgemeine demographische Daten (Alter, Geschlecht), Migrationshintergrund, sowie Motivation für das Fach Naturwissenschaften der Probanden erhoben. Über alle drei Messzeitpunkte hinweg wurden das Fachwissen zu den durchgeführten Experimenten ermittelt und jeweils ein Protokoll angefertigt (vgl. Tabelle 12). Während der Interventionszeitpunkte wurden begleitend Protokolle in der Experimentalgruppe angefertigt. Die Schule stellte die Noten des Halbjahreszeugnisses im Januar 2014 aller Schüler der Fächer Naturwissenschaften und Deutsch zur Verfügung.

Tabelle 12: Einsatz der Erhebungsinstrumente zu den jeweiligen Untersuchungszeitpunkten.

Zeitpunkt	Zu erfassendes Variable	Erhebungsinstrument
Pre-Test	Demografische Daten	Abfrage im Testheft
	Migrationshintergrund	Abfrage im Testheft
	Motivation am Fach	Items zur Motivation
	Fachwissen Experimente	Entwickelter Fachtest
	Protokoll	Protokollieritem
Begleitend	Protokolle (nur EG)	Forscherhefte
	Gruppenzugehörigkeit	Forscherhefte
Post-Test	Motivation am Fach	Items zur Motivation
	Fachwissen Experimente	Entwickelter Fachtest
	Protokoll	Protokollieritem
Follow-Up-Test	Fachwissen Experimente	Entwickelter Fachtest
	Protokoll	Protokollieritem

Der Fachtest war das Instrument zur Erhebung der abhängigen Variable Lernerfolg. Die unabhängige Variable von Forschungsfrage 2.1 (Gruppenzugehörigkeit) wurde nicht zu den Messzeitpunkten erhoben, sondern über die Identifikationsnummer während der Intervention erfasst. Die Zugehörigkeit zu einer Kleingruppe wurde ebenso als designbedingte, unabhängige Variable durch die Forscherhefte erhoben. Die Protokollqualität (unabhängige Variable von Forschungsfrage 2.2) wurde im Anschluss an den Interventionszeitraum bestimmt.

Die weiteren Erhebungsinstrumente dienten zur Erfassung der hier verwendeten Kontrollvariablen (Demographische Daten, Migrationshintergrund, Motivation am Fach). Weitere Kontrollvariablen waren die Noten im Fach Naturwissenschaften und Deutsch.

Testinstrument zur Motivation am Fach

Insgesamt wurden acht Items zur Ermittlung der allgemeinen Motivation am Fach Naturwissenschaften herangezogen. Als Antwortformat diente eine sechsstufige Likert-Skala. Das Testinstrument zur allgemeinen fachbezogenen Motivation am Fach setzt sich aus Items zum Fachinteresse (drei Items), zur Herausforderung (drei Items) und zum naturwissenschaftlichen Fähigkeitsselbstkonzept (zwei Items) zusammen. Die drei Konstrukte korrelieren miteinander und dienen zur Abdeckung der Kovariable „Motivation“, welche bei dem Besuch von Schülerlaboren oft eine Rolle spielt. Die Items wurden aus den Skalen Interesse und Herausforderung des Fragebogens zur aktuellen Motivation adaptiert (RHEINBERG, VOLLMEYER & BURNS, 2001). Das domänenspezifische Fähigkeitsselbstkonzept (zwei Items) wurde nach KUNTER ET AL. (2002) adaptiert.

Multiple Choice Fachtest

Die fachlichen Inhalte der Intervention wurden durch die Inhalte der Experimente festgelegt. Zu jedem der Experimente wurden spezifische Fachfragen für einen lernzielbezogenen Test entwickelt. Ein allgemeiner Vorwissenstest über spezielle Basiskonzepte oder chemische Grundprinzipien wurde ausgeschlossen, da die Lernziele der Experimente sehr spezifisch sind und das Vorwissen auch in dieser Hinsicht erhoben werden sollte. Es war nicht zu erwarten, dass die Schüler in diesem Bereich Vorwissen mitbringen, jedoch galt es dies durch den lernzielbezogenen Pre-Test auszuschließen. Der Fachtest bestand in der Hauptstudie aus 28 Items zu sechs Experimenten. Die Anzahl der Items lag bei vier bis sechs Items pro Experiment, welche zu einem Skalenmittelwert innerhalb eines Experimentes oder als gesamten Skalenmittelwert zusammengefasst werden. Die unterschiedliche Anzahl der Items lag an den teilweise ausführlicheren Inhalten der Experimente und den daraus resultierenden möglichen Fragen. Als Antwortformat wurde bei 24 Items Multiple Choice mit vier Antwortmöglichkeiten gewählt, bei denen nur eine Antwortmöglichkeit (Attraktor) richtig war. Bei vier Items wurde das offene Antwortformat gewählt, da die Distraktoren in diesen Fällen zu einfach waren. Die Anzahl der Fragen waren zu den Messzeitpunkten identisch. Die Reihenfolge der Items und deren Antwortmöglichkeiten wurden zu den Messzeitpunkten variiert, um

einen Wiedererkennungswert zu senken und Testwiederholungseffekte zu vermeiden (THEYßEN, 2014). Es gab zu jedem Messzeitpunkt zwei Varianten der Testhefte, um ein Abschreiben vom Sitznachbarn zu umgehen.

Die Durchführungsobjektivität muss während der Erhebungssituation gewährleistet sein, indem für alle Testpersonen vergleichbare Testbedingungen bei der Datenerhebung herrschen (RIESE & REINHOLD, 2014). Die Tests führten deshalb immer die gleiche Person im Klassenverband mit einem Ablaufplan durch. Der laufende Schulbetrieb lies es jedoch nicht zu, dass alle Klassen immer zu einem ähnlichen Zeitpunkt im Tagesverlauf getestet werden. Zur Sicherung der Auswertungsobjektivität wurde vorab ein Erwartungshorizont mit Bewertungsschlüssel in Form einer Kodieranleitung erstellt. Auf eine zweite Eingabe der Daten von einem weiteren Beurteiler wurde verzichtet.

Zur Kontrolle der Reliabilität wurde die Berechnung von Cronbachs Alpha für Pre- und Post-Test herangezogen. Ein hoher Reliabilitätswert gibt Auskunft darüber, dass eine hohe interne Konsistenz im Test herrscht. Dies bedeutet, dass alle Items sehr genau das gleiche Konstrukt messen. Cronbachs Alpha liegt für Pre- und im Post-Test zusammen bei 0.71. Nach LIENERT & RAATZ (1998) sind Reliabilitätskoeffizienten bei Gruppenvergleichen zwischen .50 und .70 als ausreichend zu betrachten.

Zur Angabe der Itemschwierigkeit von Multiple-Choice Fragen wird die Häufigkeitsverteilung der Antwortmöglichkeiten herangezogen. Nach RIESE & REINHOLD (2014) liegt die mittlere relative Lösungshäufigkeit zwischen 20% und 80%. Die Häufigkeitsverteilung der Antwortmöglichkeiten im Fachtest von Pre- und Posttest lagen insgesamt zwischen 21% bis 77% und liegen somit in einem annehmbaren Bereich für die Schwierigkeit der Fragen.

Die Trennschärfe oder Itemtestkorrelation gibt an, wie hoch ein Item mit den anderen Items der Skala korreliert. Liegt eine niedrige Korrelation zwischen den Items im Gesamtergebnis des Tests vor, haben die Items wenig miteinander gemeinsam, was darauf hindeuten kann, dass unterschiedlich latente Variablen gemessen werden. Die Trennschärfe des Gesamttests zeigt zudem, ob ein Test verschiedene Facetten eines Konstrukts abbildet und so zwischen den Probanden zu differenzieren vermag. Die Trennschärfe berechnet sich bei intervallskalierten Daten über eine Produkt-Moment-Korrelation. Für eine mittlere Trennschärfe sind Werte von $r_{it} > .30$ erstrebenswert (BORTZ & DÖRING, 2009). Die Trennschärfe des Fachwissens-testes des Pre- und Posttests liegt bei $r_{it} = .23$.

Grundsätzlich kann bei selbst entwickelten Fachtests die Validität überprüft werden, indem zum Beispiel die Ergebnisse des Fachtests mit den Ergebnissen eines Intelligenztests als Konvergenzvalidität der Konstruktvalidität korreliert werden (RIESE & REINHOLD, 2014). Da jedoch bei der Intervention aus zeitlichen Gründen nicht die Möglichkeit bestand, zusätzliche Tests (wie zum Beispiel den KFT-Test) durchzuführen, kann diese Überprüfung nicht stattfinden. Da jedoch spezifisches Fachwissen über die Experimente überprüft wurde, ist es anzunehmen, dass auch dieses Wissen geprüft wurde, da bei der Formulierung der Items auf die Inhaltsvalidität geachtet wurde.

Protokollieritem³

Die Konzeption dieses Items beruhte auf der Idee, etwa zwei Minuten lange standardisierte Videosequenzen eines Experimentes als Grundlage für das Erstellen von Erarbeitungsprotokollen in Testsituationen zu nutzen. In Einklang mit BERTHOLD ET AL. (2004) wurde angenommen, dass Videosequenzen ein geeigneter Ausgangspunkt für das Testen der Fertigkeit zum Erstellen von Erarbeitungsprotokollen sind. Die Probanden wurden durch einen schriftlichen Arbeitsauftrag dazu aufgefordert, eine Videosequenz genau zu betrachten und die präsentierten Inhalte in Form eines Erarbeitungsprotokolls darzustellen. Die Kombination aus Videosequenz und Arbeitsauftrag wird als „Video-Item“ bezeichnet. Für die Sichtung des Videos und das Erstellen eines Protokolls standen insgesamt sieben Minuten Bearbeitungszeit zur Verfügung. Im Arbeitsauftrag kann der Protokollant selbst oder eine andere Person als Adressat vorgesehen werden. Der Einsatz dieses Instrumentes diente zur Entwicklung einer Auswertungsmethode für Protokolle.

Protokolle

Während der Durchführung der sechs Experimente fertigte die Experimentalgruppe pro Proband jeweils ein Protokoll an. Das Protokoll konnte während der Experimentierphase erstellt und zusätzlich in einer fünfminütigen Protokollphase nochmal überprüft beziehungsweise überarbeitet werden. Die Anweisung für die Schüler zum Erstellen dieser Protokolle lautete *„Zeichne und Schreibe deine Ergebnisse so auf, dass du sie anderen Kindern erklären*

³ Mit Erlaubnis von Springer Nature: ENGL, L., SCHUMACHER, S., SITTER, K., GRÖBLER, M., NIEHAUS, E., RASCH, R., ROTH, J. & RISCH, B. (2015). Entwicklung eines Messinstrumentes zur Erfassung der Protokollierfähigkeit—initiiert durch Video-Items. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1),223-229.

kannst!“. Dazu stand den Schülern für jedes Experiment eine Seite zur Verfügung. Insgesamt wurden in dem Interventionszeitraum sechs mal 60 Protokolle angefertigt. In Abhängigkeit vom Adressaten kann sich das Erarbeitungsprotokoll durchaus in der Ausführlichkeit und im Detailgrad unterscheiden. Bezogen auf das Erzeugen einzelner Darstellungen konnte bereits gezeigt werden (Cox, 1999), dass Darstellungen, die für Dritte erzeugt werden, oft mit mehr Details (zum Beispiel Beschriftungen) versehen werden, als Darstellungen für den eigenen Gebrauch. Deshalb wurde in diesem Fall im Arbeitsauftrag für den Protokollanten eine andere Person als Adressat vorgesehen, um ein möglichst detailreiches Protokoll zu erhalten. Es lässt sich jedoch nicht ausschließen, dass ein Protokollant Teilbereiche eines Experimentes nicht wiedergibt, da sie für ihn selbstverständlich sind. Über die Frage, welche Teilbereiche eines Protokolls ein Protokollant notiert und welche er zurückhält, gibt es leider keine Erkenntnisse. Deshalb kann hier nicht ausgeschlossen werden, dass ein Protokollant nicht alle Inhalte des Experimentes aufschreibt, obwohl sie ihm aufgefallen sind.

3.3.3 Konzeption der Lerneinheit

Diese Teilstudie umfasste drei Doppelstunden zu je 90 Minuten. In jeder Doppelstunde wurden zwei Experimente à 40 Minuten durchgeführt, so dass insgesamt sechs Experimente in der Einheit bearbeitet wurden. Die Einheit wurde durchweg strukturgleich konzipiert, um eine Vergleichbarkeit zwischen den einzelnen Experimenten zu gewährleisten. Jede Kleingruppe erhielt zu jedem Versuch eine Kiste mit Arbeitsheft (Anhang VI) sowie alle nötigen Materialien für das Experiment. Um die Vergleichbarkeit der Ergebnisse zu gewährleisten, wurden die Experimente so konzipiert dass sie selbstständig durchgeführt werden konnten. Ansonsten lassen sich die gewonnenen Ergebnisse der Gruppen nicht miteinander vergleichen. Um Probleme bei der Benennung von Geräten zu umgehen, befand sich im Arbeitsheft eine Seite mit Bildern der Geräte und deren Namen. Bei jedem Experiment wurde im Plenum durch die Präsentation eines Bildes oder eines Filmes die Fragestellung eröffnet. Somit sollte gewährleistet werden, dass alle Schüler die Fragestellung des Experimentes nicht überlesen. Die Fragestellung wurde jedoch zusätzlich im Arbeitsheft nochmal wiederholt. Für Probleme mit dem Material oder der Handhabung beim Experimentieren stand jeder Gruppe ein Betreuer zur Seite, welcher jedoch inhaltlich keine Hilfestellung geben sollte. Nach Ablauf der ersten 40 Minuten wurde die Gruppe dazu angehalten, das Experiment zu beenden und die Materialien wieder in der Kiste zu verstauen. Anschließend wurde die nächste Fragestellung durch den Betreuer präsentiert. Falls eine Gruppe vor Ablauf der Zeit mit einem Versuch fertig war, sollten Teilaspekte des Experimentes wiederholt werden.

Die inhaltliche Konzeption der Einheit orientierte sich am Rahmenlehrplan des Fachs Naturwissenschaften für die 5. und 6. Klasse des Landes Rheinland-Pfalz. Der im Schuljahr 2009/10 eingeführte Lehrplan umfasst acht Themenfelder und baut auf den drei Grundpfeilern der naturwissenschaftlichen Grundbildung den Kompetenzen, Fachwissen und Kontexten auf (MINISTERIUM FÜR BILDUNG, WISSENSCHAFT, JUGEND UND KULTUR, 2010). Der naturwissenschaftliche Erkenntnisweg stand in der beschriebenen Unterrichtseinheit im Fokus und zielte somit auf den Kompetenzbereich Erkenntnisgewinnung ab. Themenfeld 7 des Lehrplanes legt den Schwerpunkt auf dem Kompetenzbereich Erkenntnisgewinnung und die Weiterentwicklung naturwissenschaftlicher Erkenntnismethoden (MINISTERIUM FÜR BILDUNG, WISSENSCHAFT, JUGEND UND KULTUR, 2010). Schüler sollen laut Lehrplan Stoffe durch selbst entwickelte Kriterien und spezifischen Stoffeigenschaften ordnen und Kriterien geleitet experimentelle Untersuchungen an Stoffen durchführen. Dieses Projekt wurde deshalb in der 6. Klasse als Einstieg pas-

send zu Themenfeld 7 konzipiert. Tabelle 13 gibt einen Überblick über die Experimente und deren Inhalt.

Tabelle 13: Übersicht über die eingesetzten Experimente im Verlauf der Intervention.

Doppelstunde	Experimentes	Titel	Inhalt
1	1	Was sprudelt denn da?	Die Inhaltsstoffe der Brause auf ihre Eigenschaften überprüfen.
	2	Das Gas der Brause	Das Gas der Brause über seine Eigenschaften identifizieren.
2	3	Unbekannte Brausetablette	Vergleich von Brausetabletten mit Hilfe der unterschiedlichen Masse an Kohlenstoffdioxid in Wasser.
	4	Wieviel sprudelt denn da?	Einwirkung der Wassereigenschaften auf das Löslichkeitsverhalten von Kohlenstoffdioxid.
3	5	Unbekanntes Pulver	Ein unbekanntes Pulver mit Hilfe seiner Eigenschaften identifizieren.
	6	Einfluss der Gaszusammensetzung auf die Atmosphärentemperatur	Mit Hilfe eines Modellexperimentes den Einfluss von Gasen auf den Treibhauseffekt identifizieren.

Die Inhalte der Experimente bauen aufeinander auf, so dass die Reihenfolge eingehalten werden musste. So entsteht zum Beispiel bei Experiment 1 ein Gas, welches bei Experiment 2 identifiziert wird. Bei Experiment 2 werden ebenso weitere Gase eingeführt, welche sich in Experiment 6 wiederholen. Auch die experimentellen Arbeitsweisen wurden durch die Experimente wiederholt geübt und weiter spezifiziert. So nahm das Einleiten eines Gases zum Beispiel in Experiment 3 noch einen Großteil des Experimentes ein und wurde geübt und in Experiment 6 ist das Einleiten nur ein Teilschritt im Experiment.

Experiment 1: Was sprudelt denn da?

Mit Experiment 1 wurden die Inhaltsstoffe der Brause auf ihre Eigenschaften überprüft. Den

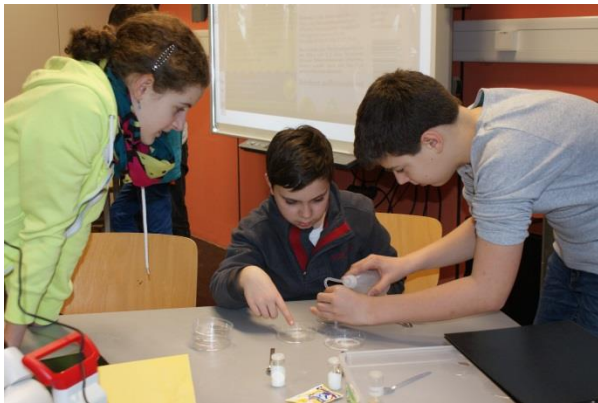


Abbildung 33: Schülergruppe prüft die Inhaltsstoffe der Brause auf ihre Eigenschaften.

Schülern wurde eine Zutatenliste der Brauseverpackung präsentiert, welche die Inhaltsstoffe Weinsäure, Zucker, Natriumhydrogencarbonat und Lebensmittelfarbe anzeigt. Den Schülern war die charakteristische Eigenschaft der Brause beim Einfüllen in Wasser bekannt. Die Fragestellung für dieses Experiment lautete deshalb „Welche Stoffe sind für das Sprudeln der Brause in Wasser verantwortlich und welche nicht?“.

Den Schülern standen zur Beantwortung der Frage ein Päckchen Brausepulver, ein Binokular, separat abgefüllte Inhaltsstoffe der Brause und diverse Arbeitsgeräte zur Verfügung. Um nochmals die Fragestellung zu verdeutlichen, beobachteten die Schüler das Brausepulver unter dem Binokular und stellten fest, dass dieses unterschiedliche Stoffe enthält. Eine Hilfefkarte zur Bedienung des Binokulars lag bei, da nicht vorausgesetzt werden kann, dass alle Schüler sich mit diesem Gerät auskennen. Die Versuchsanleitung gab vor, dass zunächst die Stoffe einzeln auf die Eigenschaft „sprudelt mit Wasser“ getestet werden sollten und dann unterschiedliche Kombinationen der Inhaltsstoffe überprüft werden sollten (vgl. Abbildung 33). Die Schüler kamen zu dem Ergebnis, dass die Inhaltstoffe Weinsäure und Natriumhydrogencarbonat in Kombination sprudeln und die Stoffe Zucker und Lebensmittelfarbe nicht sprudeln. Die Begründung der Reaktion von Carbonaten mit Säuren wurde den Schülern nicht geliefert. Das Experiment blieb auf der Ebene des Ordnen von Stoffen anhand von Kriterien. Experiment 1 verdeutlicht bereits das Grundprinzip von Themenfeld 7 des Lehrplans Naturwissenschaften: Die Stoffeigenschaften wurden identifiziert und über das Ordnen und Kombinieren der Stoffe kommen die Schüler durch experimentelle Schritte zur Schlussfolgerung einer Fragestellung.

Experiment 2: Das Gas der Brause

Experiment 2 griff den Kontext und das Ergebnis aus Experiment 1 nochmals auf. Es sollte das Gas identifiziert werden, welches entsteht, wenn Brause in Wasser gegeben wird. Zunächst wurden die drei Aggregatzustände besprochen, um sicherzustellen, dass allen Schülern der Begriff „Gas“ bekannt war. Anschließend wurden Steckbriefe der Gase Kohlenstoffdioxid, Luft und Distickstoffoxid⁴ mit



Abbildung 34: Schülergruppe beim Einleiten des unbekanntes Gases in Bromthymolblau.

ihren Eigenschaften präsentiert. Die Steckbriefe befanden sich während des Experimentes im Arbeitsheft. Mit Hilfe der Steckbriefe und zwei Teilerperimenten konnte die Fragestellung „Welches Gas entsteht wenn man Weinsäure und Natriumhydrogencarbonat mischt und welche Eigenschaften hat es?“ beantwortet werden. Die Schüler gewannen das zunächst noch unbekanntes Gas durch die Zugabe von Wasser auf ein Gemisch aus Natriumhydrogencarbonat und Weinsäure und fingen es in einem Luftballon auf. Anschließend wurde das Gas auf Brennbarkeit und Verhalten in Bromthymolblau überprüft. Das Gas löschte die Kerzen und Bromthymolblau färbte sich nach dem Einleiten des Gases gelb. Durch einen Abgleich mit den Steckbrief fanden die Schüler heraus, dass es sich um Kohlenstoffdioxid handelt. Distickstoffoxid und Luft zeigten kein saures Verhalten in Bromthymolblau und löschten auch keine Kerzenflammen. Experiment 2 baute auf der Erkenntnis des vorangegangenen Experimentes auf und festigte diese. Im Sinne des Lehrplans wurden Stoffeigenschaften experimentell erschlossen und mit vorgegebenen Daten verglichen (MINISTERIUM FÜR BILDUNG, WISSENSCHAFT, JUGEND UND KULTUR, 2010). Auch hier stand die Klassifizierung von Stoffen durch das Ordnen und Vergleichen im Vordergrund und nicht die Erläuterung der Stoffeigenschaften auf Teilchenebene. Zusätzlich schulte dieses Experiment auch das gemeinsame Arbeiten, da manche Schritte nur durch die Zusammenarbeit der Gruppenmitglieder gelöst werden konnte (vgl. Abbildung 34).

⁴ Der Begriff „Distickstoffoxid“ wurde als vereinfachter Begriff für den Stoff „Distickstoffmonooxid“ verwendet. Der Trivialname „Lachgas“ wurde aus didaktischen Gründen vermieden.

Experiment 3: Unbekannte Brausetablette

In diesem Experiment wurden drei Brausetabletten mit Hilfe der unterschiedlich entstehenden Volumina an Kohlenstoffdioxid beim Lösen in Wasser miteinander verglichen. Den Schülern wurde zunächst das Löslichkeitsverhalten von Kohlenstoffdioxid in Wasser erläutert. Ihnen war bereits aus Experiment 2 bekannt, dass beim Sprudeln der Brausetablette in Wasser Kohlenstoffdioxid entsteht. Im Wasser löst sich Kohlenstoffdioxid bis das Wasser im Messzylinder gesättigt ist. Ist das Wasser im Messzylinder mit Kohlenstoffdioxid gesättigt, drückt das weiter entstehende Kohlenstoffdioxid das Wasser aus dem Messzylinder (siehe Abbildung 35). Den Schülern wurde in einer Präsentation unterschiedliche Packungen von

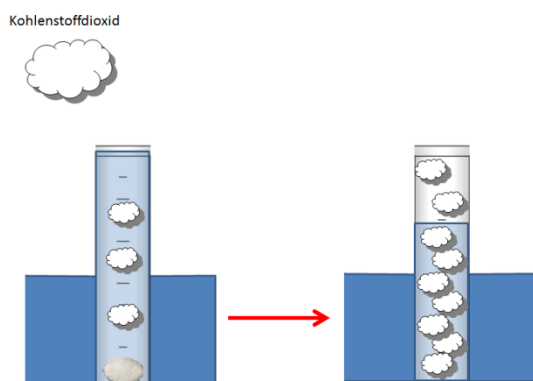


Abbildung 35: Löslichkeitsverhalten von Kohlenstoffdioxid, Ausschnitt aus der Präsentation zur Einführung.

ist die unbekannte Brausetablette?“ zu beantworten. Den Schülern standen dazu eine Wasserwanne, ein Messzylinder, die drei bekannten Brausetabletten als auch die unbekannte Brausetablette zur Verfügung. Eine Brausetablette wurde unter einen mit Wasser gefüllten Messzylinder gelegt und das Löslichkeitsverhalten beobachtet. Es war darauf zu achten, nach jeder Brausetablette das Wasser zu wechseln, da das bereits entstandene Kohlenstoffdioxid Einfluss auf die Löslichkeit weiterer Brausetabletten hat. Anschließend sollten die Schüler die Werte des entstandenen Kohlenstoffdioxids im Messzylinder nach Sättigung der bekannten Brausetabletten mit dem Wert der unbekanntten Brausetablette vergleichen. Die Brausetabletten wurden so ausgewählt, dass eindeutig vergleichbare Werte entstehen. Bei der unbekanntten Brausetablette handelte es sich um JA VITAMIN C. Auch hier stand durch Vergleichen von Stoffeigenschaften das Ziel der Aufgabenstellung zu erreichen im Fokus. Das komplexe Lösungsverhalten von Kohlenstoffdioxid in Wasser auf Teilchenebene war dabei nur ergänzend als Erläuterung der Stoffmerkmale anzusehen.

Brausetabletten (JA VITAMIN C, JA MAGNESIUM, FIT und VITAL VITAMIN C) gezeigt. Zusätzlich wurde ihnen mitgeteilt, dass sich in jeder Kiste eine unbekannte Brausetablette befindet, welche eine der bekannten Brausetabletten ist. Es galt die Fragestellung „Wie kann man unterschiedliche Brausetabletten identifizieren und welche

Experiment 4: Wieviel sprudelt denn da?

Experiment 4 thematisierte die Einwirkung der Wassereigenschaften auf das Löslichkeitsverhalten von Kohlenstoffdioxid. Die experimentelle Methode aus Experiment 3 wurde nochmal aufgegriffen und das Löslichkeitsverhalten von Kohlenstoffdioxid in Wasser wiederholt.

Den Schülern war vor der Fragestellung deutlich zu machen, welche Messwerte viel beziehungsweise wenig gelöstes Kohlenstoffdioxid anzeigen: Wenig Kohlenstoffdioxid im Messzylinder bedeutet, dass sich mehr Kohlenstoffdioxid in Wasser lösen kann und umgekehrt viel Kohlenstoffdioxid bedeutet, dass sich wenig Kohlenstoffdioxid in Wasser lösen kann (vgl. Abbildung 36). Es galt die Fragestellung zu beantworten „Unter welchen Bedingungen entsteht wie viel Kohlenstoffdioxid und wann kann mehr oder weniger Kohlenstoffdioxid gelöst werden?“. Den Schülern standen dazu eine Wasserwanne, ein Messzylinder, mehrere identische Brausetabletten, Essig und ein Wasserkocher zur Verfügung. Die Bedeutungen der Messwerte aus Abbildung 36 waren während des Experimentes an die Wand projiziert, damit keine Fehler bei der Interpretation der Ergebnisse entstehen konnten. Die Schüler veränderten nacheinander den Zustand des Wassers und verglichen dann die Versuchsbedingungen niedrige Temperatur, saures Milieu und hohe Temperatur miteinander, indem jeweils die gleiche Brausetablettensorte benutzt wurde. Nach jeder Versuchsbedingung wurde das Wasser gewechselt. In saurem Milieu entstand wenig Kohlenstoffdioxid im Messzylinder, bei niedriger Temperatur etwas mehr und bei hoher Temperatur ließ sich am meisten Kohlenstoffdioxid im Messzylinder messen. Dies bedeutet im sauren Milieu löst sich mehr Kohlenstoffdioxid im Wasser als bei niedriger Temperatur und bei hoher Temperatur löst sich am wenigsten Kohlenstoffdioxid. Ziel dieser Aufgabe war es, das methodische Vorgehen aus Experiment 3 zu wiederholen und auch den inhaltlichen Aspekt weiter zu vertiefen. Nicht die Begründung, warum löst sich mehr oder weniger Kohlenstoffdioxid stand im Fokus, sondern die vergleichende experimentelle Vorgehensweise.

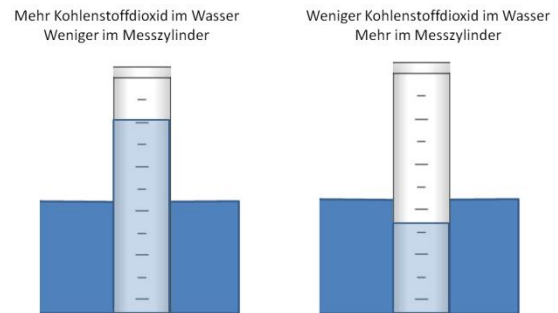


Abbildung 36: Bedeutung der Messwerte im Messzylinder beim Löslichkeitsverhalten von Kohlenstoffdioxid.

Experiment 5: Unbekanntes Pulver



Abbildung 37: Schülergruppe beim Untersuchen des unbekanntes Pulvers.

In diesem Experiment galt es die Eigenschaften eines unbekanntes Pulvers zu identifizieren und mit den Eigenschaften der bekannten Substanzen Taubenzucker, Weinsäure und Kochsalz zu vergleichen. Bei dem unbekanntes Pulver handelte es sich um eines der drei Alltagsstoffe. Den Schülern wurden Bilder von Taubenzucker, Weinsäure, Kochsalz und des unbekanntes Pulvers präsentiert. Die Frage-

stellung lautete „Wie kann man unterschiedliche Pulver identifizieren und welches ist das unbekanntes Pulver?“. Durch die Teilerperimente „Verhalten in Bromthymolblau“ und „Verhalten beim Erhitzen“ galt es die Eigenschaften der bekannten Pulver und die Eigenschaften des unbekanntes Pulvers herauszufinden (vgl. Abbildung 37). Den Schülern standen die Stoffe, das unbekanntes Pulver vermischt mit Blumenerde und alle weiteren Arbeitsmaterialien zur Verfügung. Das unbekanntes Pulver zeigte beim Erhitzen eine Verkohlung auf und Bromthymolblau blieb bei Zugabe der Substanz blau. Diese Eigenschaften konnten nur bei Traubenzucker nachgewiesen werden. Kochsalz veränderte seine Konsistenz beim Erhitzen über einer Kerze nicht und Weinsäure ließ den Indikator nach gelb umschlagen. Durch das Identifizieren der experimentell erschlossenen Stoffeigenschaften und den Vergleich der Stoffklassen wurde die Fragestellung beantwortet. Die Untersuchung von weißen Pulvern und den Vergleich der Stoffeigenschaften ist eine häufige Unterrichtsmethode in Themenfeld 7 sowie im Anfangsunterricht Chemie.

Experiment 6: Einfluss der Gaszusammensetzung auf die Atmosphärentemperatur

Die Schüler untersuchten in diesem Experiment die Einwirkung von Gasen auf die Temperatur. Ein kurzer Film über die Einwirkung von Gasen auf die Temperatur der Atmosphäre und Entstehung des Treibhauseffektes leitete das Thema ein. Den Schülern wurde dadurch verständlich gemacht, dass sich in unserer Atmosphäre außer dem Gasgemisch Luft auch noch andere Gase, unter anderem auch Kohlenstoffdioxid, befinden. Die Zusammensetzung der Gase in der Atmosphäre beeinflusst wiederum die Temperaturentwicklung auf der Erde. Es

galt durch ein Experiment die Frage zu beantworten „Wie wirken sich verschiedene Gase auf den Treibhauseffekt aus?“. Die Schüler bestrahlten die Gase Kohlenstoffdioxid und Luft über eine definierte Zeit in einem Kolben mit einer Lampe und bestimmten dabei minütlich die Temperatur (Abbildung 38). Vorab war bei beiden Durchläufen auf die Ausgangstemperatur



Abbildung 38: Schülergruppe beim Ablesen der Temperatur der Atmosphäre im Kolben.

des Gases zu achten. Nach bereits fünf Minuten zeigte sich, dass die Temperatur der Kohlenstoffdioxid-Atmosphäre, um circa 1-2°C höher angestiegen war, als die Luft-Atmosphäre. Da bei diesem Experiment vor allem Probleme bei dem Aufbau des Versuches festgestellt wurden, unterstützte eine Dauerschleife mit Hilfen zum Aufbau die Schüler fort weg beim Experimentieren. Den Schülern bereits bekannte experimentelle Methoden, wie zum Beispiel Einleiten eines Gases in ein Gefäß, waren in der Vorbereitung für dieses Experiment notwendig. Zusätzlich musste hier genau gearbeitet werden, weshalb sich der Einsatz dieses Experimentes nur mit etwas Vorerfahrung empfiehlt. Modellcharakter und Deutung des Experimentes blieben wie zuvor unberücksichtigt. Nur das vergleichende Experimentieren und die Schlussfolgerung zur gegebenen Fragestellung waren für die Untersuchung entscheidend.

3.3.4 Methoden der Datenauswertung

Die Transkription der Daten von Testheften und Protokollen wurde mit Hilfe von Kodierplänen in MICROSOFT EXCEL eingegeben und anschließend mit dem Statistikprogramm R.STUDIO VERSION 0.98.1091 ausgewertet. Zur Beantwortung der Hypothesen wurde sowohl deskriptive Statistik als auch Inferenzstatistik angewendet. Im Bereich der analytischen Statistik wurden sowohl Unterschiedsprüfungen als auch Zusammenhangsprüfungen vorgenommen. Die Begründung der Auswertungsschritte inklusive der einzelnen Verfahren wird in diesem Kapitel beleuchtet. Für alle Auswertungen gilt das gleiche Signifikanzniveau.

Das Signifikanzniveau α des p-Wertes wird nach FISCHER (1935) für alle weiteren Verfahren wie folgt festgelegt:

$p < 0.05$ = signifikant *, $p < 0.01$ = hochsignifikant **, $p < 0.001$ = höchstsignifikant *** (EID ET AL., 2011). Die Angabe der Effektgrößen für messwiederholte Varianzanalysen erfolgt durch das generalisierte η^2 (η_G^2) und bei nicht messwiederholten Varianzanalysen der Effektgröße f nach COHEN. Die Effektgröße f nach COHEN wird wie folgt zugeteilt: $f \approx 0.14$ = klein, $f \approx 0.35$ = mittel, $f \approx 0.57$ = groß (EID ET AL., 2011). Die Interpretation der Ergebnisse von η_G^2 wird wie folgt beansprucht: $\eta_G^2 \approx 0.02$ = klein, $\eta_G^2 \approx 0.13$ = mittel, $\eta_G^2 \approx 0.26$ = groß (BAKEMAN, 2005).

Überprüfung der Stichprobe

Die Teilnahme der Probanden bei den Testterminen war für die Untersuchung entscheidend, um eine Entwicklung der Lernwirksamkeit aufzeigen zu können. Deshalb wurden Probanden, welche bei den Testterminen fehlten, für die nachfolgende Auswertung nicht berücksichtigt. Aufgrund von Fehlterminen bei den Pre-Tests und Post-Tests entfielen insgesamt 11 Personen aus der Stichprobe. Auswertungen, die sich auf den Pre-Post-Vergleich beziehen, wurden deshalb mit $n = 104$ durchgeführt. Bei der Follow-Up-Erhebung fehlten weitere sieben Personen. Auswertungen, welche sich auf den Vergleich von allen drei Messzeitpunkten beziehen, wurden deshalb mit einer Stichprobe von $n = 97$ durchgeführt. Deshalb schwanken die Stichprobengrößen zwischen 97 und 104 Teilnehmer.

Vorab wurde die Vergleichbarkeit der Interventionsgruppen untersucht. Dies ist notwendig, da die Einteilung der Gruppen nicht parallelisiert vorgenommen wurde, sondern quasi-randomisiert durch die Lehrkräfte. Die Zusammensetzung der Interventionsgruppe wurde auf Unterschiede in den Kontrollvariablen Motivation, Geschlecht, Nawi-Note, Deutschnote

und Vorwissen in einzelnen Experimenten überprüft. Alle statistischen Tests beruhen auf einer bestimmten Verteilung der Stichprobe. Für die Auswahl der statistischen Tests (t-Test, ANOVA, lineare Regression) sind vor allem Normalverteilung und Varianzhomogenität entscheidend. Deshalb war vor der Prüfung auf Unterschiede bei jeder Variable eine Bestimmung auf diese Kriterien entscheidend. Die Normalverteilung wurde durch den Shapiro-Wilk-Test überprüft. Nach ROYSTON (1995) liegt bei einem p -Wert > 0.1 im Shapiro-Wilk-Test eine Normalverteilung vor. Die Annahme auf Varianzhomogenität besagt, dass die Varianzen der abhängigen Variable in beiden Populationen gleich verteilt sein muss. Die Annahme der Varianzhomogenität zeigt, dass Varianzen der abhängigen Variable in beiden Populationen gleich sind (LUHMANN, 2011). Der Levene-Test prüft die Nullhypothese, dass die Varianzen in beiden Gruppen gleich verteilt sind (ebd.). Ein signifikantes Ergebnis bedeutet also, dass die Varianzen ungleich verteilt sind und damit die Voraussetzung Varianzhomogenität gegeben verletzt ist (ebd.). Weisen die Kontrollvariablen beide Voraussetzungen auf, können die Unterschiede in der Gruppenzusammensetzung mit Hilfe einer ANOVA getestet werden. Ansonsten müssen alternative Tests für Varianzheterogenität und nicht normalverteilte Daten herangezogen werden.

Lernwirksamkeit der Einheit

Das Wissen über die Experimente wurde durch den Fachtest ermittelt. Jede richtig beantwortete Frage zählt ein Punkt, falsch beantwortete Fragen null Punkte. Die Lernwirksamkeit wird über zwei Wege definiert: Der allgemeine Lernerfolg wird durch den Vergleich der richtig beantworteten Fragen im Fachtest über die drei Messzeitpunkte ermittelt. Im Anschluss wurden die Mittelwerte der einzelnen Messzeitpunkte miteinander verglichen, um den allgemeinen Wissenszuwachs über die einzelnen Experimente aufzuzeigen. Zum anderen kann die Lernwirksamkeit auch als Lernzuwachs angegeben werden. Der Lernzuwachs ergibt sich durch die Differenz der richtigen Antworten im Pre-Test und der richtig beantworteten Fragen im Post-Test. Dieser Wert gibt Auskunft über den Zuwachs an Wissen.

Bevor ein Vergleich in den Interventionsgruppen vorgenommen wird, muss zunächst die allgemeine Lernwirksamkeit der Einheit überprüft werden. Nur ohne den vorherigen Gruppenvergleich, kann der Wissenszuwachs durch die Experimente veranschaulicht oder eventuelle Probleme mit der Einheit identifiziert werden. Mit der Voraussetzung Varianzhomogenität und Normalverteilung werden die Unterschiede der Mittelwerte Pre- zu Post-Test und

Pre- zu Follow-Up-Test mit einem t-Test für abhängige Stichproben überprüft. Der t-Test vergleicht die Unterschiede von zwei Mittelwerten, also metrischen Daten, miteinander. Mit dem t-Test wird die Nullhypothese überprüft, dass beide Stichproben aus Populationen mit identischen Mittelwerten der untersuchenden Variablen stammen. Hier wird der t-Test für abhängige Stichproben eingesetzt, da die zu vergleichenden Werte aus messwiederholten Variablen stammen.

Aufklärung von Kontrollvariablen

Um die Lernwirksamkeit ausschließlich auf die Zugehörigkeit zu der Kontroll- beziehungsweise Experimentalgruppe beziehen zu können, wurde im Vorfeld der Einfluss weiterer Variablen (Kontrollvariablen) auf die Lernwirksamkeit ausgeschlossen. Die Aufklärung der Kontrollvariablen erfolgt über beide möglichen Definitionen des Lernerfolges.

Kontrollvariablen Lernerfolg der Messzeitpunkte

Die Stichprobe lag bei $n=97$, da alle drei Messzeitpunkte miteinander verglichen wurden. Um zu prüfen, welche unabhängigen Variablen Einfluss auf den Lernerfolg nehmen, wurde eine mehrfaktorielle gemischte Varianzanalyse durchgeführt. Diese spezielle ANOVA ist vor allem für Analysen geeignet, die einen nicht-messwiederholten Faktor und mindestens einen messwiederholten Faktor (LUHMANN, 2011) haben. Hier handelt es sich bei den unabhängigen Kontrollvariablen (Geschlecht, Migrationshintergrund, Nawi-Note und Kleingruppe) um nicht metrische und nicht-messwiederholte Daten und bei der abhängigen Variablen Lernerfolg um metrische, messwiederholte Daten. Wie bei allen mehrfaktoriellen Varianzanalysen wird für jeden Faktor den Haupteffekt (hier sowohl Between- als auch Within-Effekte) sowie die Interaktionseffekte zwischen den Faktoren untersucht (LUHMANN, 2011). Bei der unabhängigen Variable „Motivation“ lagen metrische Daten vor, weshalb die Überprüfung dieser Kontrollvariablen mit einer linearen Regression überprüft wurde. In der bivariaten linearen Regression wird der lineare Zusammenhang zwischen einer metrischen unabhängigen Variablen und einer metrischen abhängigen Variablen modelliert (LUHMANN, 2011).

Kontrollvariablen Lernzuwachs

Diese Überprüfung der Kontrollvariablen bezieht sich auf den Lernzuwachs von Pre-zu Post-Zeitpunkt der sechs Experimente. Die Stichprobe umfasste deshalb auch $n=104$. Eine einfaktorielle Varianzanalyse ohne Messwiederholung diente zur Überprüfung. Werden mehr als

zwei Mittelwerte miteinander verglichen, ist der t-Test nicht mehr geeignet, da mit einem großen α -Fehler zu rechnen ist. Die einfaktorielle Varianzanalyse ohne Messwiederholung dient dem Vergleich der Mittelwerte aus einer unabhängigen Stichprobe (LUHMANN, 2011). Diese Varianzanalyse ohne Messwiederholung erklärt die Unterschiede im Lernerfolg zwischen Probanden (Between-Faktor) durch die Kontrollvariablen. Als abhängige Variable wurde jeweils der Lernzuwachs der sechs Experimente und als unabhängige Variable die Kontrollvariablen definiert. Die Varianzanalyse mit der unabhängigen Variable Motivation wurde wiederum mit einer linearen Regression überprüft. Bei allen Ergebnissen der Varianzanalysen wurden zusätzlich Effektgrößen berechnet. Um Effekte über unterschiedliche Untersuchungen hinweg miteinander vergleichen zu können, wurde von COHEN (1962) eine Effektgröße f postuliert, welche sich an der Standardabweichung in der Stichprobe relativiert.

Vergleich des Lernerfolgs der Interventionsgruppen

Diese Auswertung liefert die Hauptargumente zur Beantwortung von Forschungsfrage 2.1 und deren Hypothesen. Der Vergleich des Lernerfolgs in der Hinsicht auf die Interventionsgruppe wird ebenso über den Lernerfolg über die drei Messzeitpunkte und über den Lernzuwachs durchgeführt. Weiterhin sollen zunächst die einzelnen Experimente gegenübergestellt werden und anschließend ein Gesamtvergleich über alle sechs Experimente stattfinden. Der Lernerfolg beziehungsweise der Lernzuwachs der Experimente wird in Abhängigkeit zur zentralen unabhängigen Variable der Untersuchung (Einteilung der Probanden in Kontroll- und Experimentalgruppe) dargestellt und anschließend die Unterschiede analysiert. Für den Vergleich des Lernerfolgs zu den drei Messzeitpunkten, werden die Mittelwerte der richtig beantworteten Fragen berechnet und die Voraussetzung Normalverteilung und Varianzhomogenität überprüft. Bei der Varianzanalyse für die Mittelwerte der Fachfragen ist eine mehrfaktorielle gemischte Varianzanalyse nötig. Der messwiederholte Faktor und die abhängige Variable ist der Lernerfolg. Der nicht-messwiederholte Faktor und die unabhängige Variable ist die Gruppenzugehörigkeit zur Kontroll- beziehungsweise Experimentalgruppe. Der Lernzuwachs wird nach zwei Möglichkeiten berechnet: Der Vergleich von Pre- zu Post-Test und der Vergleich von Post- zu Follow-Up-Test. Hier ist eine anschließende Varianzanalyse ohne Messwiederholung möglich. Als abhängige Variable wurde jeweils der Lernzuwachs der sechs Experimente (einzeln oder gesamt) und als unabhängige Variable die

Gruppenzugehörigkeit bestimmt. Es soll dadurch der Unterschied im Lernerfolg zwischen den Probanden durch die Gruppenzugehörigkeit aufgeklärt werden.

Kategoriensystem zur Bewertung von Protokollen⁵

Um die Ausprägung der Qualität von Erarbeitungsprotokollen ermitteln zu können, müssen diese durch ein identisches Auswertungsverfahren beurteilt werden. Es konnte auf kein bestehendes Kategoriensystem aufgebaut werden. Deshalb wurde ein Kategoriensystem mit den Protokollen des Protokollieritems aus Teilstudie 2 (Kapitel 3.2) entwickelt. Durch die Zusammenarbeit im Forscherverbund während dieser Studie, konnte eine Vielzahl an Protokollen aus unterschiedlichen Fächern ausgewertet und zur Entwicklung der Kategorien herangezogen werden. Durch eine induktive Vorgehensweise wurden Erkenntnisse aus den einzelnen Protokollen verallgemeinert, um die Auswertungsmethode auf weitere Protokolle übertragen zu können. Die Überprüfung des Kategoriensystems auf die Gütekriterien Objektivität, Reliabilität und Validität wird in Kapitel 3.3.5 dargestellt.

Als geeignete Entwicklungsstrategie für ein Auswertungsverfahren erwies sich die qualitative Inhaltsanalyse nach MAYRING (2008). Dabei wurden aus dem Datenmaterial Kategorien abgeleitet, unter denen möglichst viele Daten zusammengefasst werden können. Die Kategorien werden entweder induktiv aus dem Material gewonnen oder deduktiv (theoriegeleitet) an das Material herangetragen (MAYRING, 2008). Wir entschieden uns hier für eine Mischform, bei der zunächst induktiv Kategorien aus dem Material gewonnen und auf der Grundlage vorhandener theoretischer Erkenntnisse ergänzt und verfeinert wurden. Natürlich lässt sich bei einer induktiven Vorgehensweise schon erarbeitetes theoretisches Wissen nicht restlos ausblenden, sodass möglicherweise implizit bereits erste Kriterien vorlagen. Diese Kritik wird im Zusammenhang mit der qualitativen Inhaltsanalyse immer wieder vorgebracht. Hierbei gilt jedoch zu beachten, dass dieses theoretische Wissen den Prozess der Identifizierung wichtiger Aspekte in den vorliegenden Materialien unterstützt. Den Anfang des beschriebenen qualitativ-inhaltsanalytischen Vorgehens bildete die Sichtung der Erarbeitungsprotokolle aus Teilstudie 2. In mehreren Arbeitsschritten wurden erste Kategorien formuliert, die das Konstrukt der Protokollierfähigkeit genauer beschreiben. Dazu wurden zunächst je zwei

⁵ Mit Erlaubnis von Springer Nature: ENGL, L., SCHUMACHER, S., SITTER, K., GRÖßLER, M., NIEHAUS, E., RASCH, R., ROTH, J. & RISCH, B. (2015). Entwicklung eines Messinstrumentes zur Erfassung der Protokollierfähigkeit–initiiert durch Video-Items. *Zeitschrift für Didaktik der Naturwissenschaften*. 21(1), 223-229.

Erarbeitungsprotokolle herausgegriffen, die sich deutlich voneinander unterschieden. Besonderheiten und Auffälligkeiten wurden in Form zweier möglichst gegensätzlicher Merkmalspaare (zum Beispiel „sinnvoller Aufbau“ vs. „eher willkürliche Aneinanderreihung der Inhalte“) formuliert (RÜEDE & WEBER, 2012). Die so entstandenen gegensätzlichen Merkmalspaare wurden im Anschluss nach Ähnlichkeit sortiert und Kategorien in Anlehnung an einschlägige Literatur (zum Beispiel LANGER, SCHULZ VON THUN & TAUSCH, 1981; BECKER-MROTZEK & BÖTTCHER, 2011; BRUNER, 1971) gebildet. Auf diese Weise entstanden vier Kategorien.

Die Kategorie „Gliederung“ wurde an das Hamburger Verständlichkeitsmodell der Psychologen LANGER, SCHULZ VON THUN & TAUSCH (1981) angelehnt. Auf der Suche nach Eigenschaften, die einen Text verständlich machen, definieren die genannten Autoren vier Verständlichkeitsdimensionen, darunter auch das Merkmal „Gliederung“, das zunächst sowohl die äußere als auch innere Struktur des Textes berücksichtigt. Die äußere Struktur erfasst rein optische Merkmale des Textes wie zum Beispiel die Verwendung von Absätzen oder Strukturierungselementen, wie Aufzählungszeichen, Pfeile, Beschriftungen oder ähnliches. Es geht also um eine Strukturierung, die zunächst nicht auf den Inhalt bezogen ist. Bei der inneren Struktur geht es um Fragen wie: Stehen die Sätze beziehungshaltig nebeneinander und werden sie folgerichtig aufeinander bezogen? Ist die gewählte Reihenfolge der Darstellung sinnvoll (LANGER ET AL., 1981)? Unter der Kategorie „Gliederung“ fassen wir lediglich die äußere Form der Darstellung. Die innere Struktur der Erarbeitungsprotokolle fließt hingegen in die Kategorie „Beziehungen“ mit ein. Im Zentrum von Protokollen stehen immer die Inhalte der Videos des Protokollieritems zu den Pre-, Post-, und Follow-Up-Messzeitpunkten. Die inhaltliche Dimension wird häufig bei der Bewertung von Schülertexten im Deutschunterricht genutzt (zum Beispiel BECKER-MROTZEK & BÖTTCHER, 2011; BÖTTCHER & BECKER-MROTZEK, 2009). BÖTTCHER & BECKER-MROTZEK (2009) zählen beispielsweise den Inhalt zu den wichtigen Basisdimensionen zur Bewertung von Schülertexten. In der Kategorie „Produkt“ wird geprüft, ob die wesentlichen Inhalte des Videos erfasst wurden und die Kategorie „Beziehungen“ prüft, ob einzelne Inhaltsbereiche beziehungshaltig oder isoliert voneinander dargestellt sind. Als weitere Kategorie wurden die Repräsentationsformen aufgenommen. Nach BRUNER (1971) lassen sich drei Formen der Repräsentation von Wissen unterscheiden: die enaktive (durch Handlungen), die ikonische (durch Bilder) und die symbolische (durch Zeichen, Sprache). Dabei hat jede Repräsentationsform ihre eigenen Vor- und Nachteile. So kann zum Beispiel ein Beziehungsgefüge, das symbolisch repräsentiert ist, in ein räumliches Bild (ein Vorstel-

lungsbild) übersetzt werden, um die gestellte Aufgabe besser und direkter verarbeiten zu können. Durch die Simultaneität eines Bildes werden zeitliche Segmente, welche den Ablauf einer Handlung steuern, mitunter weniger deutlich. Gewisse Zusammenhänge können nur in Worten leicht und treffend ausgedrückt werden (SCHNOTZ, 1994). Aus diesen Gründen ist es nicht sinnvoll objektive Gütekriterien anzugeben, die für eine bestimmte Repräsentationsform sprechen. Deshalb wird die Kategorie „Repräsentationsform“ rein deskriptiv erfasst und nicht normativ bewertet. Die Kategorien „Gliederung“, „Produkt“ und „Beziehungen“ unterliegen normativen Auswertungskriterien. Tabelle 14 veranschaulicht die abgeleiteten Kategorien und den Auswertungsmöglichkeiten.

Tabelle 14: Kategorien und beinhaltete Aspekte der Auswertung.

Kategorie	Aspekte
Gliederung	<ul style="list-style-type: none"> • Absätze • Strukturierungselemente • Lesbarkeit
Produkt	<ul style="list-style-type: none"> • Korrekte Darstellung wesentlicher Inhalte (Anzahl je nach Inhalt)
Beziehungen	<ul style="list-style-type: none"> • Korrekte Darstellung wesentlicher Beziehungen zwischen Inhalten (Anzahl je nach Inhalt) • Innere Struktur, Systematik des Aufbau
Repräsentationsform (deskriptiv)	<ul style="list-style-type: none"> • Art der gewählten Repräsentationsform (Skizze, Tabelle, Text, Fachsymbole) • Verwendung mehrerer Repräsentationen

Die Auswertung der Kategorien erfolgte mittels Erfassung der verschiedenen Aspekte jeder Kategorie. Beispielsweise besteht die Kategorie Gliederung aus den Aspekten Absätze, Strukturierungselemente und Lesbarkeit (vgl. Tabelle 14). Jeder Aspekt wird durch ein Item repräsentiert. Die Kategorien „Repräsentationsform“ und „Gliederung“ sind unabhängig vom Inhalt und haben deshalb eine feste Anzahl an Items. Die Itemanzahl der Kategorien „Produkt“ und „Beziehungen“ ergibt sich für jede Videosequenz aus den dargestellten Inhalten. Jedes Item wird dichotom kodiert (0 = nicht vorhanden, 1 = vorhanden). Eine Kodieranleitung definiert mit Ankerbeispielen die Ausprägungen der Items (vgl. Anhang VII).

Abbildung 39 veranschaulicht das mögliche Vorgehen bei der Auswertung der Aspekte. Ein Ausschnitt aus der Kategorie Produkt zeigt neun von insgesamt 17 Items. Die Inhalte wurden in dem Protokoll erschlossen und als vorhanden oder nicht vorhanden notiert.

Das Wort „Harnstoff“ ist vorhanden	<input checked="" type="checkbox"/>
Das Wort „Calciumchlorid“ ist vorhanden.	<input checked="" type="checkbox"/>
Das Wort „Kochsalz“ ist vorhanden.	<input checked="" type="checkbox"/>
„Einwirkung der Stoffe auf die Wassertemperatur“ ist vorhanden.	<input checked="" type="checkbox"/>
Das Ereignis „1.Messzeitpunkt“ ist vorhanden.	<input checked="" type="checkbox"/>
Die Zahl „23,7“ ist vorhanden.	<input checked="" type="checkbox"/>
Die Zahl „23,6“ ist vorhanden.	<input checked="" type="checkbox"/>
Die Zahl „23,7“ ist ein zweites Mal vorhanden.	<input checked="" type="checkbox"/>
Das Ereignis „Stoffe werden eingerührt“ ist vorhanden.	<input checked="" type="checkbox"/>

Abbildung 39: Vorgehen bei der Bewertung von Protokollen. Links ein Erarbeitungsprotokoll, rechts ein Ausschnitt aus dem Kodierplan der Kategorie Produkt. Der Haken symbolisiert „vorhanden“ und das Kreuz „nicht vorhanden“ des Produkts.

Für eine gezielte Aussage über die Qualität der Erarbeitungsprotokolle war die dichotome Kodierung jedoch nicht ausreichend. Um diese Aussage zu quantifizieren, wurden die einzelnen Items nach der Kodierung gewichtet (siehe Anhang VII) und zu einem Gesamtwert für die jeweilige Kategorie aggregiert. Ein Protokoll wurde durch drei Kategorien (Gliederung, Produkt und Beziehung) bewertet und konnte in jeder Kategorie einen maximalen Wert von 1 erreichen. Mit Hilfe dieses Wertes konnten Bewertungen über die jeweilige Kategorie getroffen werden. Für die Bewertung des gesamten Protokolls wurden die drei Werte addiert, womit der höchstmögliche Wert für ein Protokoll bei 3 lag. Diese Auswertung ermöglichte allgemeine Rückschlüsse über die Güte eines Erarbeitungsprotokolls.

Überprüfung auf Zusammenhänge zwischen Lernwirksamkeit und Protokollqualität

Die Überprüfung auf einen Zusammenhang zwischen Lernwirksamkeit und Protokollqualität kann nur bei der Experimentalgruppe stattfinden, da diese Protokolle angefertigt haben. Die beiden Merkmale werden exemplarisch an den Ergebnissen der Lernwirksamkeit eines Experimentes mit dem dazu angefertigten Protokoll ausgewertet. Die Auswahl für ein Experiment erfolgt nach der Auswertung der Lernwirksamkeit. Die Stichprobe bezieht sich auf $n=51$ für alle Probanden der Experimentalgruppe in einem Experiment.

Am Kategoriensystem zur Bewertung von Protokollen werden Änderungen auf der inhaltlichen Ebene, in den Kategorien Produkt und Beziehung, vorgenommen. Im Anschluss findet eine Auswertung der Protokolle aus diesem Experiment statt. Die Ergebnisse des Lernzuwachses dieses Experimentes werden mit der Qualität der Protokolle aus diesem Experiment miteinander in Beziehung gesetzt. Um das Maß der Protokollqualität in Zusammenhang mit dem Lernerfolg in diesem Experiment zu bringen, wurde eine Korrelation berechnet. Je nach Skalenniveau müssen unterschiedliche Korrelationskoeffizienten herangezogen werden. Da in diesem Fall zwei metrische Variablen miteinander auf Zusammenhang geprüft werden, wird die Produkt-Moment-Korrelation nach PEARSON angewendet (EID ET AL., 2011). Mit Hilfe eines Korrelationskoeffizienten kann ein Wert angegeben werden, der aussagt, wie eng zwei Variablen miteinander zusammenhängen (ebd.). Die Korrelation gibt die Stärke des linearen Zusammenhangs an. Je größer der Wert, umso stärker ist der Zusammenhang. Nach COHEN (1988) lassen sich die Korrelationen der Produkt-Moment-Korrelation wie folgt interpretieren: $r_{XY} \approx 0,10$ entspricht einem schwachen Zusammenhang, $r_{XY} \approx 0,30$ einem mittleren Zusammenhang und $r_{XY} \approx 0,50$ einem starken Zusammenhang (EID ET AL., 2011).

Einflüsse auf die Protokollqualität

Eine Varianzanalyse der Protokollqualität soll zeigen, durch welche weiteren Variablen die Unterschiede der Qualität entstehen. Die Protokolle wurden während der Intervention angefertigt und unterscheiden sich nicht im Messzeitpunkt. Für die nominalskalierten unabhängigen Variablen (Geschlecht, Nawi-Note, Deutsch-Note, Migrationshintergrund, Kleingruppe) wird das Verfahren der ANOVA angewendet. Für den Vergleich der metrischen unabhängigen Variablen (Motivation, Lernerfolg) wird die Varianz durch eine bivariate lineare Regression geprüft.

3.3.5 Gütekriterien des Kategoriensystems

Das Kategoriensystem zur Bewertung von Protokollen wurde während dieser Teilstudie entwickelt. Um anhand dieser Kategorien Aussagen über die Protokolle machen zu können, muss dieses zunächst auf die Gütekriterien überprüft werden. Die Einhaltung von Objektivität, Reliabilität und Validität sind notwendig und stellt die Güte einer wissenschaftlichen Arbeit sicher. Nur durch die Gültigkeit dieser Kriterien ist ein universeller Einsatz des Kategoriensystems gewährleistet. Es werden die Vergleichbarkeit der Ergebnisse, die Genauigkeit, sowie der Untersuchungsgegenstand des Kategoriensystems untersucht und bewertet. Diese Untersuchungen fanden am Rande der Teilstudien statt und bedingen deshalb unterschiedliche Stichproben.

Objektivität

Die Objektivität stellt ein Hauptgütekriterium für wissenschaftliche Tests dar und kann folgendermaßen definiert werden:

Das Hauptgütekriterium Objektivität kann als erfüllt gelten, wenn ein psychologischer Test hinsichtlich seiner Handhabung, Auswertung und Interpretation so festgelegt ist, dass sein Einsatz unabhängig von umgebenden Faktoren (wie Ort, Zeit, Testleiter und Auswerter) vorgenommen werden kann und unterschiedliche Testauswerter bei den gleichen Personen immer zu gleichen Ergebnissen kommen (POSPECHILL, 2010).

Die Auswertungsobjektivität durch das Kategoriensystem lässt sich durch die Übereinstimmung der Testergebnisse verschiedener Testauswerter messen und in Form einer statistischen Kennzahl über Konkordanzmaße (zum Beispiel Konkordanzkoeffizient Kendalls W, Interrater-Reliabilitätsmaß Scotts Pi) oder korrelative Maße (zum Beispiel Intraclasskorrelation) angegeben (ebd.). Es wurden einige zufällig ausgewählte Protokolle hinsichtlich der einzelnen Aspekte der vorgestellten Kategorien von mehreren geschulten Ratern unabhängig voneinander bewertet. Diese Ratings wurden anhand der Kodieranleitung vorgenommen, um anschließend die Interrater-Reliabilität zu ermitteln. Als Übereinstimmungsmaß wurde der Fleiss'Kappa-Koeffizient κ_m gewählt. Bezüglich der Interpretation geben BORTZ & DÖRING (2006) in Anlehnung an FLEISS & COHEN den Bereich zwischen 0.60 und 0.75 als zufriedenstellende Reliabilität an. Es wurden die Kappa-Koeffizienten der einzelnen Aspekte der Katego-

rien berechnet. Die drei Aspekte der Kategorie „Gliederung“, als eine fach- beziehungsweise inhaltsunabhängige Kategorie, wurden dabei über alle Teilprojekte hinweg von vier Ratern (für insgesamt 60 Protokolle) eingeschätzt. Für die einzelnen Aspekte der Kategorie „Gliederung“ liegen die κ_m -Werte zwischen 0.66 und 0.73, was nach BORTZ & DÖRING (2006) für eine gute Übereinstimmung spricht. Für die Kategorien „Produkt“ und „Beziehungen“ haben jeweils drei (fachspezifische) Rater, insgesamt rund 70 Protokolle unabhängig voneinander ausgewertet. Für die Kategorie „Produkt“ liegen die κ_m -Werte über die verschiedenen Fachinhalte (Aspekte der Kategorie) hinweg zwischen 0.67 und 0.92. Ähnliche κ_m -Werte wurden auch für die Kategorie „Beziehungen“ erzielt. Hier liegen die κ_m -Wert der relevanten Aspekte zwischen 0.74 und 0.89. Nahezu perfekt ($\kappa_m = 0.82$ bis 0.92) wurden die Aspekte der Kategorie „Repräsentationsformen“ eingeschätzt. Auf der Basis der Beurteilerübereinstimmung ist das Kategoriensystem als objektiv einzuschätzen.

Reliabilität

Die Reliabilität gibt die Zuverlässigkeit des Messinstruments und damit die Genauigkeit bezogen auf Messfehler an (MOOSBRUGGER & KELAVA, 2008). Angegeben wird die Reliabilität über den Reliabilitätskoeffizienten r , welcher einen Wert zwischen -1 und 1 annehmen kann (POSPECHILL, 2010). Liegt der Reliabilitätskoeffizient beim Wert 1, so ist das Testergebnis ohne Messfehler und ein wiederholter Testdurchlauf würde zu dem gleichen Ergebnis führen. Ein Reliabilitätskoeffizient gegen 0 hingegen bedeutet, dass das Testergebnis nur durch Messfehler ermittelt wurde und eine wiederholte Messung zu einem gänzlich anderen Ergebnis führt. Ein Wert von -1 zeigt einen negativen linearen Zusammenhang (LUDWIG-MAYERHOFER, 2012). Akzeptabel ist eine Reliabilität ab $r=0,7$ und Werte von $r=0,8-0,9$ werden als gut eingestuft (POSPECHILL, 2010). Die Reliabilität kann über die Retest-, Paralleltest-, Testhalbierungs-Reliabilität (Splithalf-Reliabilität) oder die Innere Konsistenz überprüft werden. In diesem Fall wurde die Paralleltest-Reliabilität herangezogen. Deren Vorteil liegt darin Erinnerungseffekte aufgrund der parallelen Testformen auszuschließen und dadurch die Reliabilität nicht zu unterschätzen.

Zur Bestimmung der Paralleltest-Reliabilität wurde eine Untersuchung im April 2015 mit Schülern der 5. und 6. Klasse der AG „Chemische Experimente“ (N=20) der IGS Wörth durchgeführt. Die Schüler verfassten zwei Erarbeitungsprotokolle zu zwei inhaltlich ähnlichen Video-Items zum gleichen Messzeitpunkt. Anschließend wurden beide Protokolle über

das Kategoriensystem bewertet und die Korrelation der Protokolle berechnet. Die Reliabilitätskoeffizienten sind für die einzelnen Kategorien in Tabelle 15 zusammengefasst. Für die Repräsentationsformen unterscheiden sich die Reliabilitätskoeffizienten voneinander. Die Repräsentationsform Tabelle zeigt einen negativen Zusammenhang von $r=-0.16$, während die Repräsentationsformen Text und Skizze einen positiven Zusammenhang von $r=0.68$ und $r=0.76$ zeigen. Der Reliabilitätskoeffizient der Kategorie Gliederung liegt bei $r=0.04$, für die Kategorie Beziehung $r=0.44$ und bei der Kategorie Produkt bei einem Wert von $r=0.85$. Für die Berechnung des Gesamtwertes eines Protokolls liegt der Koeffizient bei $r=0.08$.

Tabelle 15: Reliabilitätskoeffizient für zwei Videoitems.

Kategorie	Reliabilitätskoeffizient (r)
Repräsentationsform Tabelle	-0.16
Repräsentationsform Skizze	0.76
Repräsentationsform Text	0.68
Gliederung	0.04
Produkt	0.85
Beziehung	0.44
Protokoll Gesamt	0.08

Die unterschiedlichen Werte für die Kategorie Repräsentationsform lassen sich durch die unterschiedliche Nutzung von Repräsentationsformen in den Protokollen begründen. Die Schüler nutzen für ähnliche Sachverhalte nicht die gleiche Repräsentationsform. Vor allem die Repräsentationsform Tabelle wird nicht wiederholt genutzt. Auch die Kategorie Gliederung zeigt sich als keine beständige Variable in dieser Auswertungsmethode. Die beiden anderen zentralen Auswertungskategorien Produkt und Beziehung zeigen jedoch eine ganz annehmliche Reliabilität. Ein weiterer möglicher Grund wäre, dass die beiden Video-Items keine exakten Itemzwillinge sind. Dies lässt sich jedoch schwer im Vorfeld prüfen. Weiterhin muss berücksichtigt werden, dass für das zweite Video-Item ein neuer Kodierplan erstellt wurde, der noch nicht an einer anderen Stichprobe pilotiert wurde.

Validität

Die Validität gibt die Übereinstimmung an, zwischen dem was gemessen werden soll und dem was tatsächlich gemessen wird (MOOSBRUGGER & KELAVA, 2008). Mit der Kriteriumsvalidität wird die Korrelation zwischen einer Testvariable und einem gewählten Außenkriterium gegeben, das heißt durch ein Verhalten des Probanden innerhalb einer Testsituation wird auf ein Verhalten außerhalb der Testsituation geschlossen (ebd.). Diese Validitätsuntersuchung hat weiterhin den Vorteil, dass mit der Stichprobe der Reliabilitätsuntersuchung gearbeitet werden kann. Die Erarbeitungsprotokolle der Schüler wurden mit Hilfe des Kategoriensystems bewertet. Die Bewertung des Erarbeitungsprotokolls diente damit als das zu untersuchende Merkmal innerhalb der Testsituation und wurde mit einem Außenkriterium in Zusammenhang gestellt. Das Außenkriterium wäre in diesem Falle die Darstellung des gesehenen Experiments. Hierzu wurden die Schüler in einem Abstand von drei Wochen zum Tag der ersten Datenerhebung mit der Methode des Lauten Denkens befragt und rekonstruieren mithilfe ihres Protokolls den Versuch. Die Forschungsmethode des Lauten Denkens diente vor allem dazu in mentale Prozesse, wie Denk-, Lern- und Problemlöseprozesse des Schreibens von Protokollen Einblick zu erhalten (SANDMANN, 2014). Die Rekonstruktion des Experiments wurde digital aufgezeichnet und das gesprochene Wort ebenfalls mit dem Kategoriensystem bewertet. Im Sinne der Validität wird die Korrelation zwischen den Werten aus beiden Datenerhebungen in Form des Validitätskoeffizienten r_{xy} gefasst. In diesem Fall dem Testergebnis x und dem Außenkriterium y . Der Validitätskoeffizient kann einen Wert zwischen den Werten 0 und 1 annehmen. Ein Wert von 0 zeigt keinen Zusammenhang zwischen dem Testergebnis und dem Außenkriterium auf, während ein Wert von 1 eine vollständige Deckung der Merkmale repräsentiert. Die Produkt-Moment-Korrelation lässt wie folgt interpretieren: $r_{xy} \approx 0,10$ entspricht einem schwachen Zusammenhang, $r_{xy} \approx 0,30$ einem mittleren Zusammenhang und $r_{xy} \approx 0,50$ einem starken Zusammenhang (EID ET AL., 2011).

Zur Überprüfung des Kategoriensystems auf das Gütekriterium der Validität wurde der Validitätskoeffizient zwischen den Bewertungen der naturwissenschaftlichen Befragung und dem Erarbeitungsprotokoll berechnet. Die Ergebnisse sind in Tabelle 16 dargestellt. Der Validitätskoeffizient für die Repräsentationsform Tabelle liegt bei $r=0.67$, für die Skizze bei $r=0.83$ und bei Text $r=1$. Die Validitätskoeffizienten für die Kategorien Produkt, Beziehung und Protokoll Gesamt liegen sehr nah beieinander. Während er für die Kategorie Produkt bei

$r=0.91$ liegt, erreicht er für die Kategorie Protokoll Gesamt einen Wert von $r=0.92$ und für die Kategorie Beziehung einen Wert von $r=0.94$.

Tabelle 16: Validitätskoeffizient der Bewertungen der naturwissenschaftlichen Befragung und dem Erarbeitungsprotokoll.

Kategorie	Validitätskoeffizient (r_{xy})
Repräsentationsform Tabelle	0.67
Repräsentationsform Skizze	0.83
Repräsentationsform Text	1
Produkt	0.91
Beziehung	0.94
Protokoll Gesamt	0.92

Das Kategoriensystem kann durch die hohen Werte der Validitätskoeffizienten als valide eingeordnet werden. Die Kategorie Gliederung ist aus der Bewertung für die Validität ausgeschlossen, da es sich um eine deskriptive Kategorie handelt, die nicht durch die ausgewählten Methoden zur Überprüfung der Validität erhoben werden kann.

3.3.6 Untersuchungsergebnisse

Die Untersuchungsergebnisse und die anschließende Diskussion der Ergebnisse beziehen sich direkt auf die gestellten Forschungsfragen. Zunächst werden die Voraussetzungen in der Lernervariable überprüft. Im Anschluss werden die Ergebnisse der Lernwirksamkeit der Unterrichtseinheit dargestellt. Um den Lernerfolg ausschließlich auf die Zugehörigkeit zu der Kontroll- beziehungsweise Experimentalgruppe beziehen zu können, muss im Vorfeld der Einfluss weiterer Variablen (Kontrollvariablen) auf die Lernwirksamkeit ausgeschlossen werden. Die abhängige Variable Lernwirksamkeit wird zunächst auf mögliche Kontrollvariablen untersucht, um eine abschließende Varianzanalyse zu der hypothesenuntersuchenden unabhängigen Variable Gruppenzugehörigkeit vornehmen zu können. Danach erfolgt eine Überprüfung auf Zusammenhang von Lernwirksamkeit eines Experimentes und der Qualität des dazu angefertigten Protokolls.

Im Bereich der Kontrollvariable Motivation hat sich über die Messzeitpunkte keine Veränderung ergeben. Da Motivation nur als Kontrollvariable eingesetzt wurde, war dies zu erwarten. Im folgenden Ergebnissteil wird deshalb immer nur der Wert der Motivation aus dem Pre-Test verwendet.

Überprüfung der Stichprobe

Die Überprüfungen der Varianz der Variablen Motivation, Geschlecht, Vornoten und Vorwissen der Experimente ergaben keine signifikanten Unterschiede in Bezug auf die Zugehörigkeit zu den Interventionsgruppen (vgl. Tabelle 17). Es galten somit gleiche Bedingungen der Lernervariable in beiden Gruppen.

Tabelle 17: Angabe des p-Werts (inklusive F-Wert, Freiheitsgrad) und der Effektgröße f nach Cohen der Varianzanalyse einzelner Variablen in Abhängigkeit zur Zugehörigkeit zu den Interventionsgruppen. Gerundet auf zwei Nachkommastellen.

Überprüfung auf Variable	
Motivation	F(95) = 0.88, p = 0.35, f = 0.09
Geschlecht	F(95) = 0.05, p = 0.82, f = 0.02
Nawi-Note	F(95) = 2.05, p = 0.16, f = 0.13
Deutschnote	F(95) = 0.04, p = 0.84, f = 0.02
Vorwissen Experiment 1	F(95) = 1.15, p = 0.29, f = 0.11
Vorwissen Experiment 2	F(95) = 0.08, p = 0.78, f = 0.03
Vorwissen Experiment 3	F(95) = 0.04, p = 0.84, f = 0.02
Vorwissen Experiment 4	F(95) = 1.08, p = 0.31, f = 0.11
Vorwissen Experiment 5	F(95) = 0.09, p = 0.75, f = 0.03
Vorwissen Experiment 6	F(95) = 0.89, p = 0.35, f = 0.09

Wichtige Voraussetzungen für die Durchführung der Testverfahren sind die Normalverteilung und die Varianzhomogenität der eingesetzten Variablen. Alle Ergebnisse des Shapiro-Wilk-Test und des Levene-Tests werden aufgrund der Übersichtlichkeit nicht angegeben. Die Testergebnisse zeigen, dass alle p-Werte der Überprüfung auf Normalverteilung >0.1 sind. Die Ergebnisse des Levene-Tests zeigen ebenso keine signifikanten Ergebnisse. Es lagen also normalverteilte und varianzhomogene Variablen vor. Die beschriebenen Tests können eingesetzt werden.

Forschungsfrage 2.1: Welche Bedeutung hat das Protokollieren für den Lernerfolg?

(H1) Das Anfertigen eines Protokolls, während des Experimentierens, erhöht den Lernerfolg der Schüler im Vergleich zu Schülern, die kein Protokoll anfertigen.

Lernwirksamkeit der Einheit

Tabelle 18 zeigt die Mittelwerte der richtig beantworteten Fragen mit der jeweiligen Itemanzahl aller Experimente für die drei Messzeitpunkte der Stichprobe von $n=97$. Die Ergebnisse des t-Tests für die Mittelwerte der Experimente 1, 2, 3, 5 und 6 zeigen sowohl von Pre- zu Post-Test als auch von Post- zu Follow-Up-Test signifikante Ergebnisse. In diesen Experimenten und auch über den Gesamtwert der Experimente 1-6 lassen sich signifikante Unterschiede im Mittelwert über die Messzeitpunkte hinweg aufzeigen. Der Mittelwert des Lernerfolgs ist vom Follow-Up-Zeitpunkt etwas geringer als vom Post-Zeitpunkt. Bei Experiment 4 ist ein geringer Anstieg des Mittelwertes im Lernerfolg über die Messzeitpunkte hinweg zu sehen. Alle weiteren Überprüfungen wurden deshalb ohne Experiment 4 gemacht.

Tabelle 18: Lernerfolg: Auswertung des Fachtests zu den einzelnen Messzeitpunkten, angegeben in Mittelwerte und p-Werte zur Überprüfung der Unterschiede.

Experiment	Anzahl Items	Mittelwert Pre-Test	Mittelwert Post-Test	Mittelwert Follow-Up-Test	t-Test für abhängige Stichproben p-Wert Pre zu Post	t-Test für abhängige Stichproben p-Wert Post zu FU
1	4	1.82	3.12	1.3	$p < 0.001$ ***	$p < 0.001$ ***
2	7	2.73	4.12	3.72	$p < 0.001$ ***	$p < 0.001$ ***
3	5	1.54	2.63	2.61	$p < 0.001$ ***	$p < 0.001$ ***
4	4	1.05	1.41	1.00	$p = 0.006$ **	$p = 0.554$
5	4	1.22	2.05	1.87	$p < 0.001$ ***	$p < 0.001$ ***
6	4	1.31	2.56	1.98	< 0.001 ***	< 0.001 ***
Gesamtwert 1-6	28	9.82	16.07	13.88	$p < 0.001$ ***	$p < 0.001$ ***

Aufklärung von Kontrollvariablen

Kontrollvariablen Lernerfolg der Messzeitpunkte

Tabelle 19 zeigt die p-Werte der mehrfaktoriellen Varianzanalyse verschiedener unabhängiger Variablen auf den Lernerfolg der Experimente 1, 2, 3, 5 und 6. Ein p-Wert, welcher größer als das Signifikanzniveau von $\alpha=0.05$ ist, drückt aus, dass sich die abhängigen Variable Lernerfolg nicht in Abhängigkeit auf die angegebene Variable unterscheidet. Die Ergebnisse der Interaktionseffekte der Variablen Geschlecht, Migrationshintergrund, Motivation und Deutschnote zeigen im Lernerfolg der einzelnen Experimente keine signifikanten Unterschiede. Diese Variablen nahmen keinen Einfluss auf den Lernerfolg. In der Variable „Kleingruppe“ ist in Experiment 1 ein signifikanter Wert ($F(62, 130)=1.82, p=0.002^*$, $\eta_G^2=0.25$) zu verzeichnen. Die Zugehörigkeit zu einer Kleingruppe hat bei Experiment 1 demnach einen mittleren Effekt auf den Lernerfolg. Weitere signifikante Werte zeigen sich in der unabhängigen Variable Nawi-Note bei Experiment 1 ($F(8, 184)=2.63, p=0.009^{**}$, $\eta_G^2=0.05$) und bei Experiment 5 ($F(8, 184)=2.39, p=0.018^*$, $\eta_G^2=0.05$). Zusätzlich wurde der Einfluss der Kontrollvariablen noch auf die Summe des Lernerfolgs über alle sechs Experimente untersucht. Die Variable Kleingruppe zeigt einen hochsignifikanten p-Wert und einen kleinen Effekt ($F(62, 130)=1.85, p=0.0017^{**}$, $\eta_G^2=0.23$) auf den Lernerfolg aller Experimente.

Tabelle 19: Interaktionseffekte der Kontrollvariablen auf den Lernerfolg der Experimente, angegeben in p-Wert und generalisiertes eta² (η_G^2). Die Ergebnisse wurden auf zwei Nachkomma stellen gerundet.

Experiment	Ge- schlecht	Migrations- hintergrund	Kleingruppe	Motivation	Nawi-Note	Deutsch- Note
1	p = 0.32 $\eta_G^2 < 0.01$	p = 0.15 $\eta_G^2 = 0.02$	p = 0.002 * $\eta_G^2 = 0.25$	p = 0.89 $\eta_G^2 < 0.01$	p = 0.009** $\eta_G^2 = 0.05$	p = 0.21 $\eta_G^2 = 0.03$
2	p = 0.47 $\eta_G^2 < 0.01$	p = 0.54 $\eta_G^2 < 0.01$	p = 0.17 $\eta_G^2 = 0.22$	p = 0.28 $\eta_G^2 < 0.01$	p = 0.29 $\eta_G^2 = 0.03$	p = 0.78 $\eta_G^2 = 0.01$
3	p = 0.48 $\eta_G^2 < 0.01$	p = 0.93 $\eta_G^2 < 0.01$	p = 0.18 $\eta_G^2 = 0.21$	p = 0.49 $\eta_G^2 < 0.01$	p = 0.99 $\eta_G^2 < 0.01$	p = 0.96 $\eta_G^2 < 0.01$
5	p = 0.58 $\eta_G^2 < 0.01$	p = 0.93 $\eta_G^2 < 0.01$	p = 0.24 $\eta_G^2 = 0.25$	p = 0.22 $\eta_G^2 < 0.01$	p = 0.018 * $\eta_G^2 = 0.05$	p = 0.39 $\eta_G^2 = 0.03$
6	p = 0.85 $\eta_G^2 < 0.01$	p = 0.91 $\eta_G^2 < 0.01$	p = 0.18 $\eta_G^2 = 0.23$	p = 0.77 $\eta_G^2 < 0.01$	p = 0.99 $\eta_G^2 < 0.01$	p = 0.56 $\eta_G^2 = 0.02$
Gesamtwert 1-6	p = 0.39 $\eta_G^2 < 0.01$	p = 0.46 $\eta_G^2 < 0.01$	p = 0.0017 ** $\eta_G^2 = 0.23$	p = 0.19 $\eta_G^2 < 0.01$	p = 0.37 $\eta_G^2 = 0.01$	p = 0.58 $\eta_G^2 = 0.01$

Kontrollvariablen Lernzuwachs

Tabelle 20 zeigt die Ergebnisse der ANOVA für den Interaktionseffekt der Kontrollvariablen für den Lernzuwachs von Pre- zu Post-Zeitpunkt. Die Werte aus Tabelle 20 lassen darauf schließen, dass der Lernerfolg sich nicht in Abhängigkeit auf die Variablen Geschlecht, Motivation und Deutschnote unterscheidet. Bei der Variable Migrationshintergrund gibt es bei Experiment 1 einen Hinweis für eine Abhängigkeit des Lernzuwachses ($F(101)=3.71$, $p=0.03^*$, $f=0.27$), bei der Variable Kleingruppe ergibt sich ebenso ein signifikantes Ergebnis bei Experiment 1 ($F(72)=1.86$, $p=0.02^*$, $f=0.89$), sowie ein signifikantes Ergebnis bei der Variable Nawi-Note und dem Lernzuwachs bei Experiment 5 ($F(99)=2.52$, $p=0.05^*$, $f=0.32$). Der Einfluss der Kontrollvariablen wurde ebenso über den Lernerfolg über den Gesamtwert der sechs Experimente hinweg überprüft. Die Kontrollvariable Kleingruppe zeigt wieder eine Abhängigkeit auf den Lernerfolg ($F(72)=2.08$, $p\text{-Wert}<0.01^{**}$, $f=0.95$).

Tabelle 20: Interaktionseffekte der Kontrollvariablen auf den Lernzuwachs von Pre- zu Posttest der sechs Experimente. Die Ergebnisse wurden auf zwei Nachkomma stellen gerundet.

Experiment	Geschlecht	Migrationshintergrund	Kleingruppe	Motivation	Nawi-Note	Deutsch-Note
1	p = 0.41 f= 0.08	p = 0.03 * f= 0.27	p = 0.02 * f= 0.89	p = 0.17 f= 0.09	p = 0.56 f= 0.18	p = 0.51 f= 0.18
2	p = 0.28 f= 0.11	p = 0.32 f= 0.15	p = 0.13 f= 0.77	p = 0.32 f= 0.12	p = 0.46 f= 0.19	p = 0.71 f= 0.15
3	p = 0.45 f= 0.08	p = 0.79 f= 0.07	p = 0.06 f= 0.82	p = 0.27 f= 0.12	p = 0.89 f= 0.11	p = 0.71 f= 0.15
5	p = 0.38 f= 0.09	p = 0.92 f= 0.04	p = 0.48 f= 0.66	p = 0.4 f= 0.19	p = 0.05 * f= 0.32	p = 0.46 f= 0.19
6	p = 0.53 f= 0.06	p = 0.65 f= 0.09	p = 0.06 f= 0.82	p = 0.06 f= 0.03	p = 0.95 f= 0.09	p = 0.87 f= 0.11
Gesamtwert 1-6	p = 0.39 f= 0.08	p = 0.44 f= 0.13	p < 0.01 ** f= 0.95	p = 0.12 f= 0.016	p = 0.41 f= 0.2	p = 0.36 f= 0.07

Die Kontrollvariablen Geschlecht, Motivation und Deutschnote zeigen keine Abhängigkeit auf den Lernerfolg sowohl in Überprüfung der einzelnen Messzeitpunkte als auch beim Lernzuwachs. Die Variablen Migrationshintergrund, Kleingruppe und Nawi-Note beeinflussen zumindest in einzelnen Experimenten den Lernerfolg. Da im Anschluss nur der Einfluss und der Effekt der Interventionsgruppen aufgezeigt werden soll, bietet sich der Vergleich mit diesen Teilwerten deshalb nicht an. Da vor allem in Experiment 1 zweimal eine Abhängigkeit

durch andere Variablen aufgetreten ist, wird kein weiterer Vergleich anhand dieses Experiments vorgenommen.

Vergleich der Lernwirksamkeit der Interventionsgruppen

Im folgenden Abschnitt wird die Lernwirksamkeit der Experimente in Abhängigkeit zur zentralen unabhängigen Variable der Untersuchung (Einteilung der Probanden in Kontroll- und Experimentalgruppe) dargestellt und anschließend die Unterschiede analysiert. Die Lernwirksamkeit des Experimentes wird sowohl als Mittelwert der einzelnen Messzeitpunkte sowie auch als Lernzuwachs Pre- zu Post-Test und Post- zu Follow-Up-Test angegeben. Die Angaben erfolgen zunächst für die Experimente 2, 3, 5 und 6 und anschließend über den Gesamtwert aller Items des Fachtests. Die Anzahl der Stichprobe beträgt $n=97$.

Lernwirksamkeit Experiment 2

Die Anzahl der Items zu Experiment 2 beträgt sieben, sodass ein Maximum von 7 richtig beantworteten Fragen erreicht werden kann. Abbildung 40 zeigt, dass beide Gruppen beim Pre-Test einen ähnlichen Wert der richtig beantworteten Fragen aufweisen (KG $\chi=2.68$, $\sigma=1.6$, EG $\chi=2.78$, $\sigma=1.7$). Beim Post-Test erreicht die Kontrollgruppe einen Mittelwert von 3.94 ($\sigma=1.6$) und die Experimentalgruppe einen Wert von 4.33 ($\sigma=1.5$) richtig beantworteter Fragen. Ein geringerer Mittelwert ist beim Follow-Up-Test zu verzeichnen (KG $\chi=3.53$, $\sigma=1.8$, EG $\chi=3.93$, $\sigma=1.7$). Die messwiederholte ANOVA ergibt einen p-Wert von 0.69 ($F(2, 190)=0.38$, $p=0.69$, $\eta_G^2 < 0.01$).

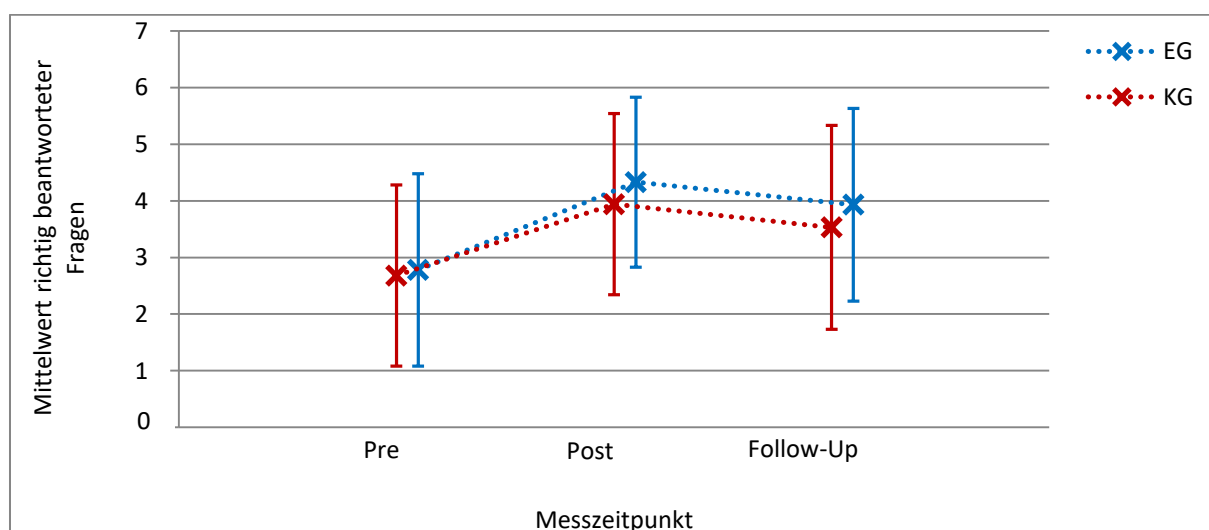


Abbildung 40: Lernerfolg: Auswertung von Experiment 2 über die drei Messzeitpunkte im Vergleich Kontroll- und Experimentalgruppe.

Abbildung 41 zeigt den Lernzuwachs Pre- zu Post-Test (links) und Post- zu Follow-Up-Test (rechts). Ein positiver Wert beim Lernzuwachs an, wie viele Fragen mehr von einem Messzeitpunkt zum nächsten richtig beantwortet wurden. Ein negativer Wert gibt entsprechend an, wie viele Fragen weniger beantwortet wurden. Von Pre- zu Post-Zeitpunkt haben beide Treatmentgruppen einen positiven Wert zu verzeichnen (KG $\chi = 1.27$, EG $\chi = 1.61$). Von Post- zu Follow-Up-Zeitpunkt verzeichnen beide Gruppen einen geringen negativen Lernzuwachs (KG $\chi = -0.41$, EG $\chi = -0.39$). Eine Varianzanalyse ohne Messwiederholung für den Lernerfolg Pre-Post ergibt einen p-Wert von 0.37 ($F(1, 102) = 0.81$, $p = 0.37$, $f = 0.09$) und bei Post- Follow-Up einen p-Wert von 0.96 ($F(1, 95) = 0.01$, $p = 0.96$, $f < 0.01$).

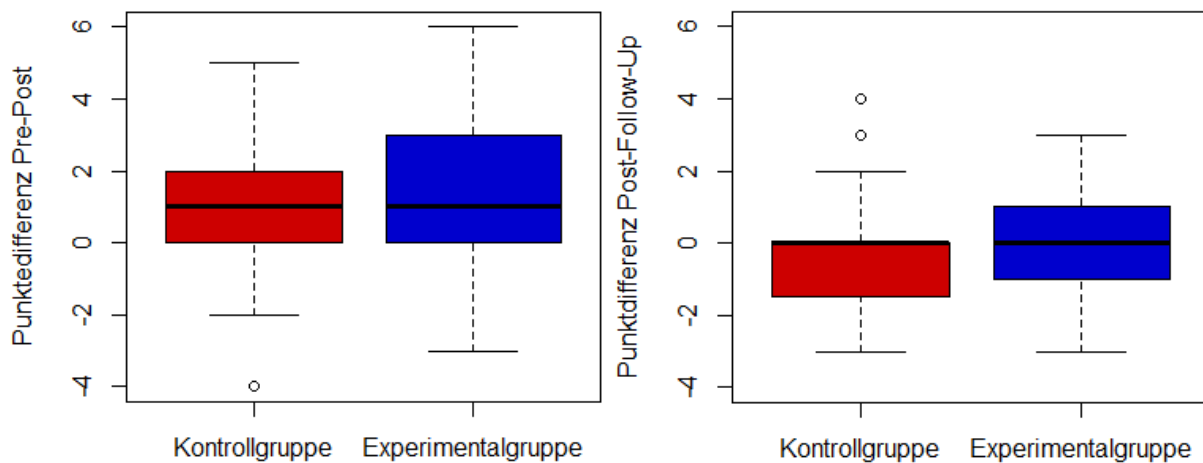


Abbildung 41: Lernzuwachs: Auswertung von Experiment 2 in einer Punktdifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.

Lernwirksamkeit Experiment 3

Die Anzahl der Items zu Experiment 3 beträgt 5, sodass ein Maximum von 5 richtig beantworteten Fragen erreicht werden kann. Abbildung 42 zeigt, dass beide Gruppen beim Pre-Test einen ähnlichen Wert der richtig beantworteten Fragen annehmen (KG $\chi = 1.61$, $\sigma = 1.09$, EG $\chi = 1.57$, $\sigma = 1.01$). Beim Post-Test erreicht die Kontrollgruppe einen Mittelwert von 2.45 ($\sigma = 1.25$) und die Experimentalgruppe einen Wert von 2.83 ($\sigma = 1.04$) richtig beantworteter Fragen. Die Experimentalgruppe weist einen geringeren Mittelwert beim Follow-Up-Test im Vergleich zum Post-Test auf, der Mittelwert der Kontrollgruppe steigt dagegen etwas an (KG $\chi = 2.57$, $\sigma = 0.9$, EG $\chi = 2.65$, $\sigma = 1.1$). Die messwiederholte ANOVA ergibt einen p-Wert von 0.23 ($F(2, 190) = 1.51$, $p = 0.23$, $\eta^2 < 0.01$).

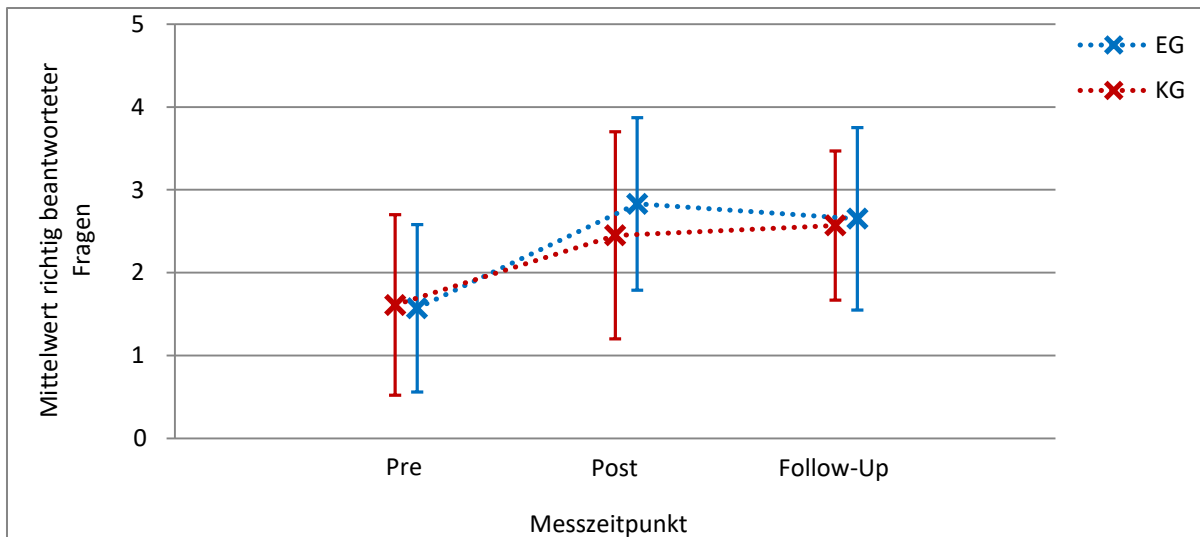


Abbildung 42: Lernerfolg: Auswertung von Experiment 3 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.

Abbildung 43 zeigt den Lernzuwachs Pre- zu Post-Test (links) und Post-zu Follow-Up-Test (rechts). Von Pre- zu Post-Zeitpunkt verzeichnen beide Treatmentgruppen einen positiven Lernzuwachs (KG $\chi = 0.85$, EG $\chi = 1.17$). Von Post- zu Follow-Up-Zeitpunkt weist die Kontrollgruppe einen geringen positiven Lernzuwachs und die Experimentalgruppe einen geringen negativen Lernzuwachs (KG $\chi = 0.12$, EG $\chi = -0.17$) auf. Eine Varianzanalyse ohne Messwiederholung für die Lernerfolg Pre-Post ergibt einen p-Wert von 0.24 ($F(1, 102) = 1.38$, $p = 0.24$, $f = 0.12$) und bei Post- Follow-Up einen p-Wert von 0.23 ($F(1, 95) = 1.48$, $p = 0.23$, $f = 0.12$).

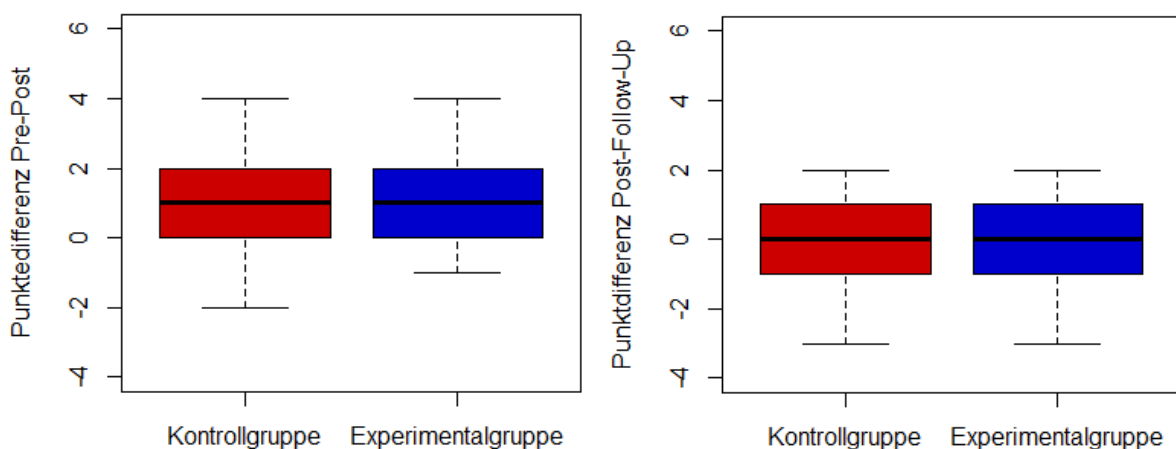


Abbildung 43: Lernzuwachs: Auswertung von Experiment 3 in einer Punktdifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.

Lernwirksamkeit Experiment 5

Die Anzahl der Items zu Experiment 5 beträgt 4, sodass ein Maximum von 4 richtig beantworteten Fragen erreicht werden kann. Abbildung 44 zeigt, dass beide Gruppen beim Pre-Test einen ähnlichen Wert der richtig beantworteten Fragen annehmen (KG $\chi = 1.27$ $\sigma = 0.8$, EG $\chi = 1.21$, $\sigma = 0.99$). Beim Post-Test erreicht die Kontrollgruppe einen Mittelwert von 2.07 ($\sigma = 1.05$) und die Experimentalgruppe einen Wert von 2.15 ($\sigma = 0.8$) richtig beantworteter Fragen. Ein geringerer Mittelwert ist beim Follow-Up-Test zu verzeichnen (KG $\chi = 1.86$, $\sigma = 1.1$, EG $\chi = 1.87$, $\sigma = 1.1$). Die messwiederholte ANOVA ergibt einen p-Wert von 0.88 ($F(2, 190) = 0.13$, $p = 0.88$, $\eta_G^2 < 0.01$).

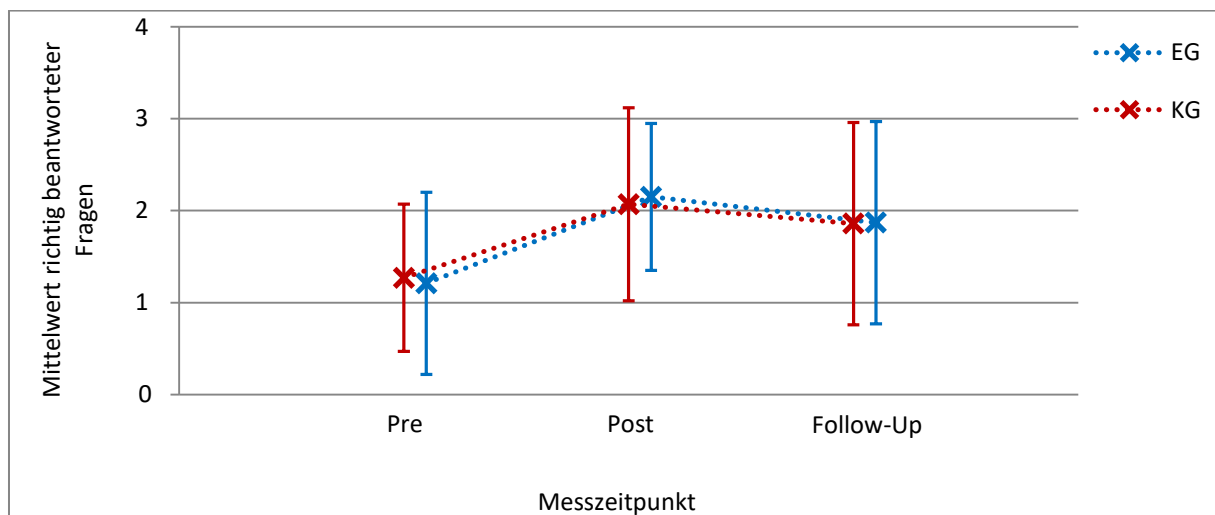


Abbildung 44: Lernerfolg: Auswertung von Experiment 5 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.

Abbildung 45 zeigt den Lernzuwachs Pre- zu Post-Test (links) und Post-zu Follow-Up-Test (rechts). Von Pre- zu Post-Zeitpunkt weisen beide Treatmentgruppen einen positiven Lernzuwachs (KG $\chi = 0.81$, EG $\chi = 0.85$) auf. Von Post- zu Follow-Up-Zeitpunkt verzeichnen beide Gruppen einen geringen negativen Lernzuwachs (KG $\chi = -0.22$, EG $\chi = -0.28$). Eine Varianzanalyse ohne Messwiederholung für den Lernerfolg Pre-Post ergibt einen p-Wert von 0.89 ($F(1, 102) = 0.02$, $p = 0.89$, $f = 0.01$) und bei Post- Follow-Up einen p-Wert von 0.79 ($F(1, 95) = 0.07$, $p = 0.79$, $f = 0.03$).

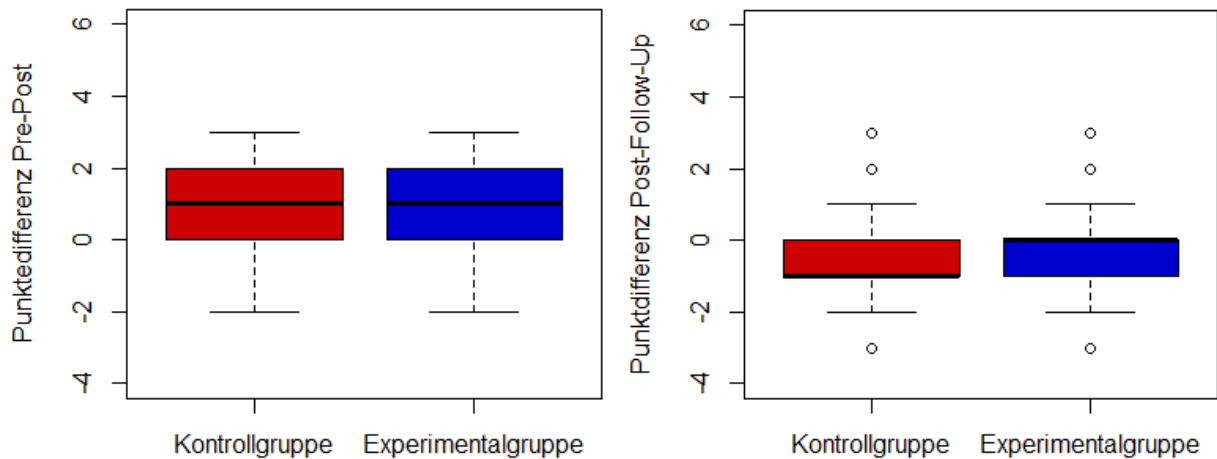


Abbildung 45: Lernzuwachs: Auswertung von Experiment 5 in einer Punktdifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.

Lernwirksamkeit Experiment 6

Die Anzahl der Items zu Experiment 6 beträgt 4, sodass ein Maximum von 4 richtig beantworteten Fragen erreicht werden kann. Abbildung 46 zeigt, dass beide Gruppen beim Pre-Test einen ähnlichen Wert der richtig beantworteten Fragen annehmen (KG 1.45 $\sigma = 1.1$, EG 1.24, $\sigma = 1.1$). Beim Post-Test erreicht die Kontrollgruppe einen Mittelwert von 2.39 ($\sigma = 1.2$) und die Experimentalgruppe einen Wert von 2.85 ($\sigma = 1.01$) richtig beantworteter Fragen. Ein geringerer Mittelwert ist beim Follow-Up-Test zu verzeichnen (KG $\chi = 1.96$, $\sigma = 1.1$, EG $\chi = 2.0$, $\sigma = 1.3$). Die messwiederholte ANOVA ergibt einen p-Wert von 0.056 ($F(2, 190) = 2.93$, $p = 0.056$, $\eta_G^2 = 0.01$).

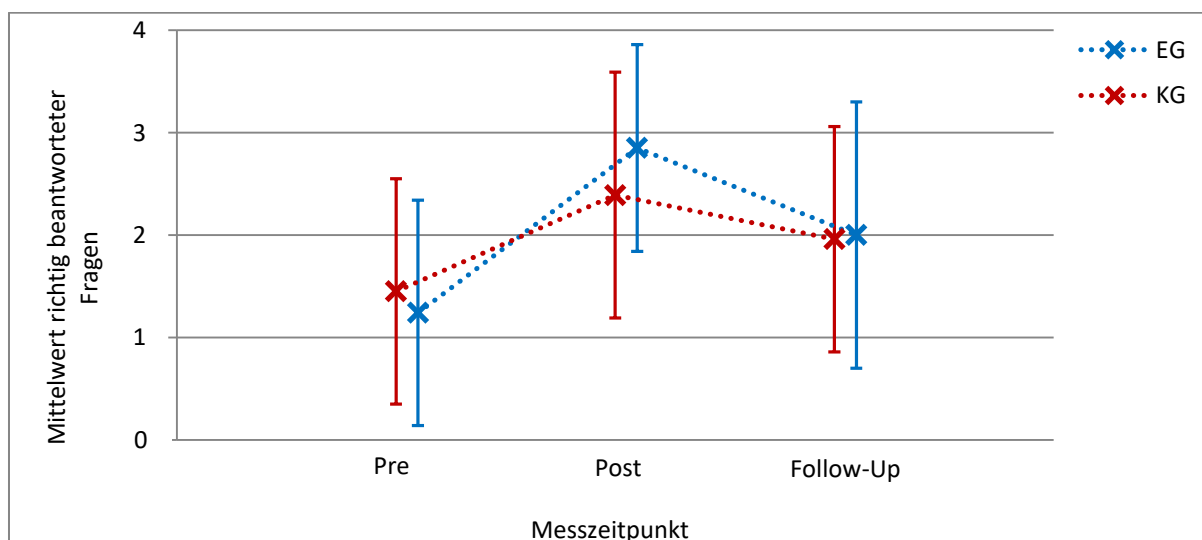


Abbildung 46: Lernerfolg: Auswertung von Experiment 6 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.

Abbildung 47 zeigt den Lernzuwachs Pre- zu Post-Test (links) und Post-zu Follow-Up-Test (rechts). Von Pre- zu Post-Zeitpunkt weisen beide Treatmentgruppen einen positiven Lernzuwachs (KG $\chi = 0.94$, EG $\chi = 1.56$) auf. Von Post- zu Follow-Up-Zeitpunkt verzeichnen beide Gruppen einen geringen negativen Lernzuwachs (KG $\chi = -0.43$, EG $\chi = -0.84$). Eine Varianzanalyse ohne Messwiederholung für den Lernerfolg Pre-Post ergibt einen signifikanten Unterschied und einen mittleren Effekt ($F(1, 102) = 6.37$, $p = 0.01^*$, $f = 0.25$) und bei Post-Follow-Up einen p-Wert von 0.17 ($F(1, 95) = 1.95$, $p = 0.17$, $f = 0.14$).

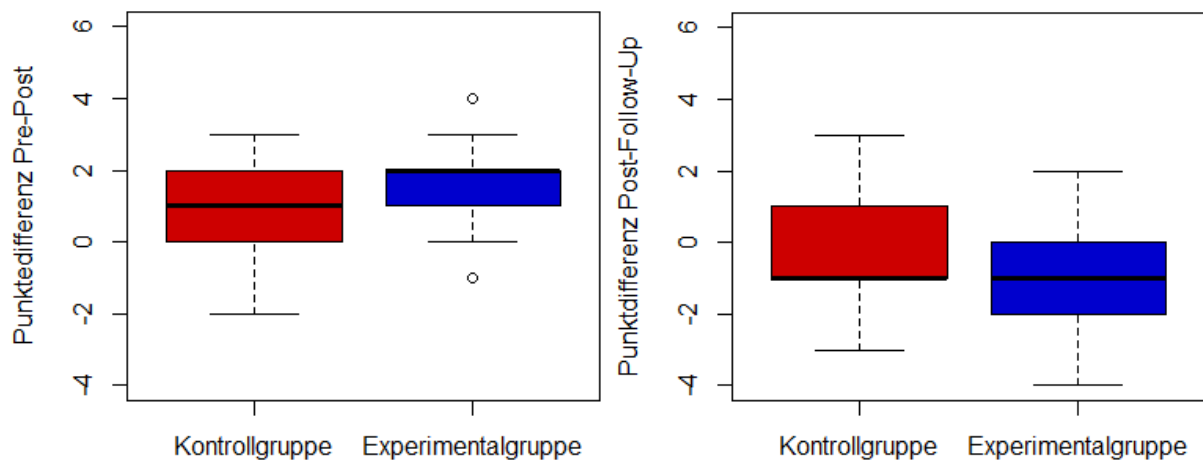


Abbildung 47: Lernzuwachs: Auswertung von Experiment 6 in einer Punktdifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.

Lernwirksamkeit der Experimente 2, 3, 5 und 6

Die Gesamtzahl der Items beträgt 21, sodass ein Maximum von 21 richtig beantworteten Fragen erreicht werden kann.

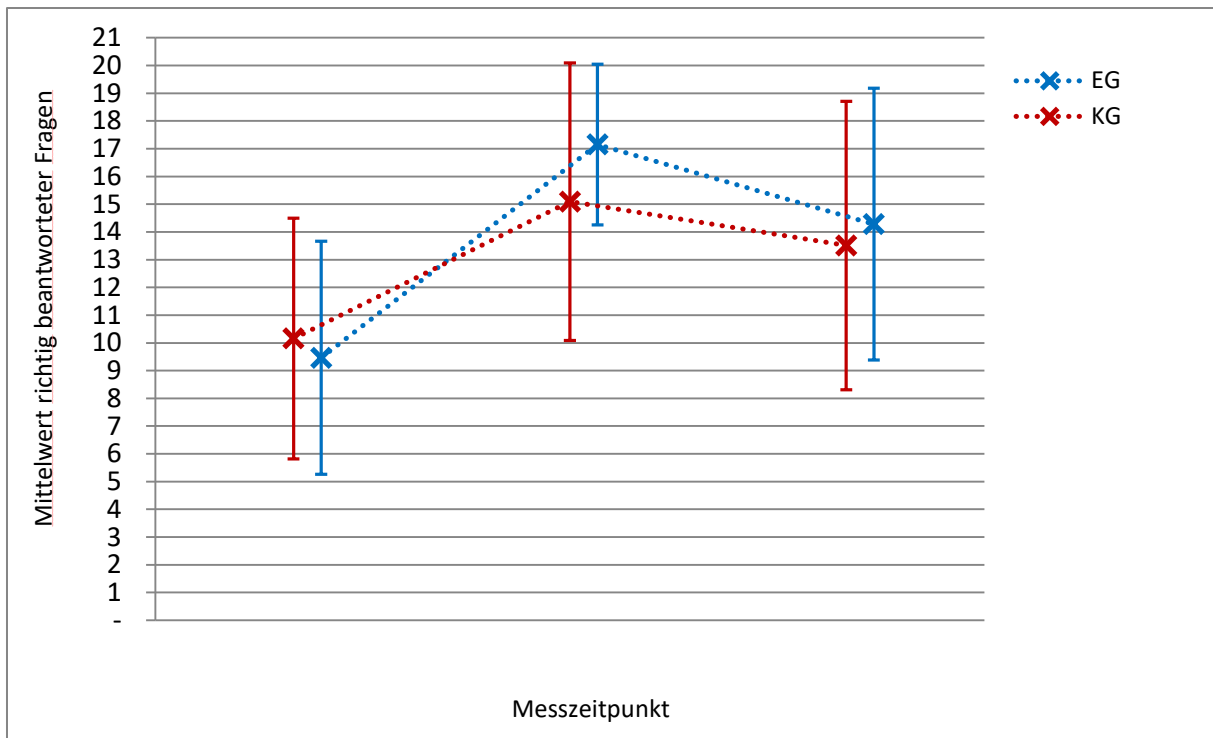


Abbildung 48 zeigt, dass beide Gruppen beim Pre-Test einen ähnlichen Wert der richtig beantworteten Fragen aufweisen (KG $\chi = 10.16$ $\sigma = 4.34$, EG $\chi = 9.46$, $\sigma = 4.2$). Beim Post-Test erreicht die Kontrollgruppe einen Mittelwert von $\chi = 15.09$ ($\sigma = 5.0$) und die Experimentalgruppe einen Wert von $\chi = 17.15$ ($\sigma = 3.9$) richtig beantworteter Fragen. Ein geringerer Mittelwert ist beim Follow-Up-Test zu verzeichnen (KG $\chi = 13.51$, $\sigma = 5.2$, EG $\chi = 14.28$, $\sigma = 4.9$). Die messwiederholte ANOVA ergibt einen Interaktionseffekt und einen kleinen Effekt ($F(2, 190) = 4.23$, $p = 0.016^*$, $\eta^2 = 0.02$). Der Post-hoc-Test (FSLD > 1,32) zeigt, dass die die Interaktion nur bei Pre-Post liegt.

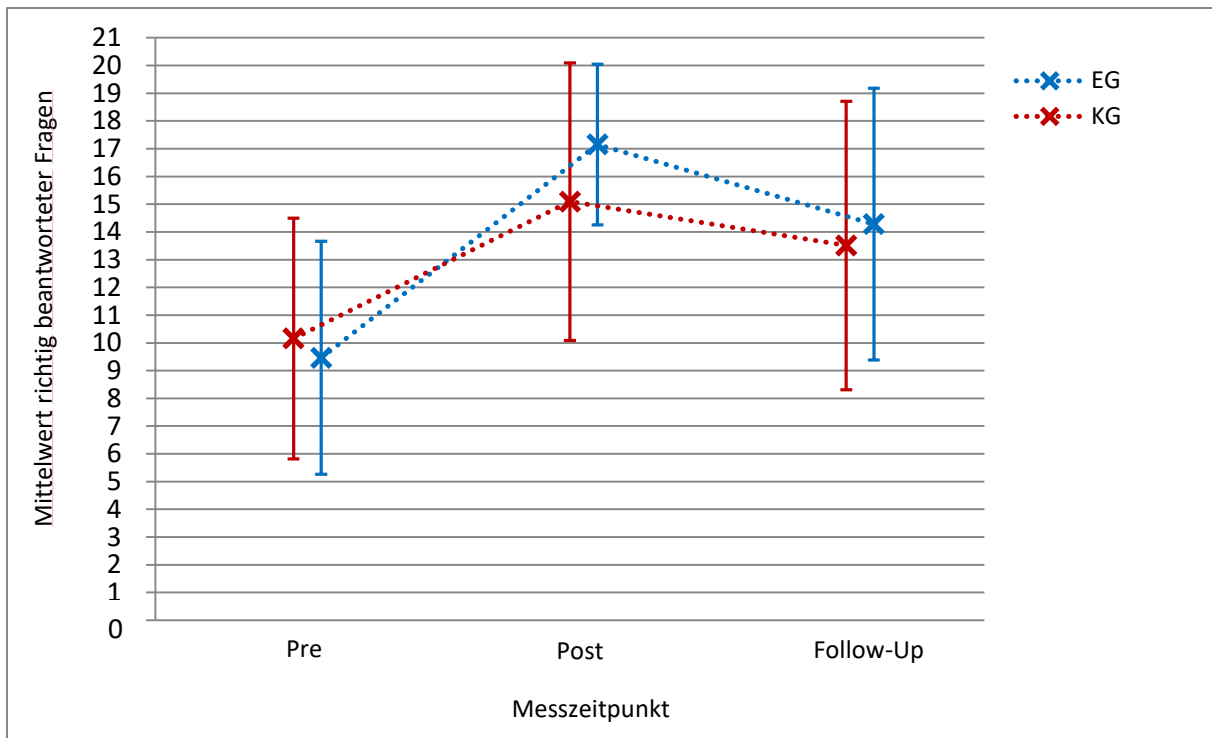


Abbildung 48: Lernerfolg: Auswertung der Experimente 2, 3, 5 und 6 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.

Abbildung 49 zeigt den Lernzuwachs Pre- zu Post-Test (links) und Post-zu Follow-Up-Test (rechts). Von Pre- zu Post-Zeitpunkt weisen beide Treatmentgruppen einen positiven Lernzuwachs (KG $\chi = 4.98$, EG $\chi = 7.41$) auf. Die Kontrollgruppe beantwortet über alle Experimente hinweg im Vergleich zum Pre-Test im Post-Test 4,98 Fragen mehr richtig. Die Experimentalgruppe beantwortet über alle Experimente hinweg im Vergleich zum Pre-Test im Post-Test 7,4 Fragen mehr richtig. Von Post- zu Follow-Up-Zeitpunkt verzeichnen beide Gruppen einen negativen Lernzuwachs (KG $\chi = -1.59$, EG $\chi = -2.87$). Eine Varianzanalyse ohne Messwiederholung für den Lernzuwachs Pre-Post ergibt einen signifikanten Unterschied und einen mittleren Effekt ($F(1, 102) = 6.83$, $p = 0.01^*$, $f = 0.26$) und bei Post-Follow-Up einen p-Wert von 0.15 ($F(1, 95) = 2.01$, $p = 0.15$, $f = 0.15$).

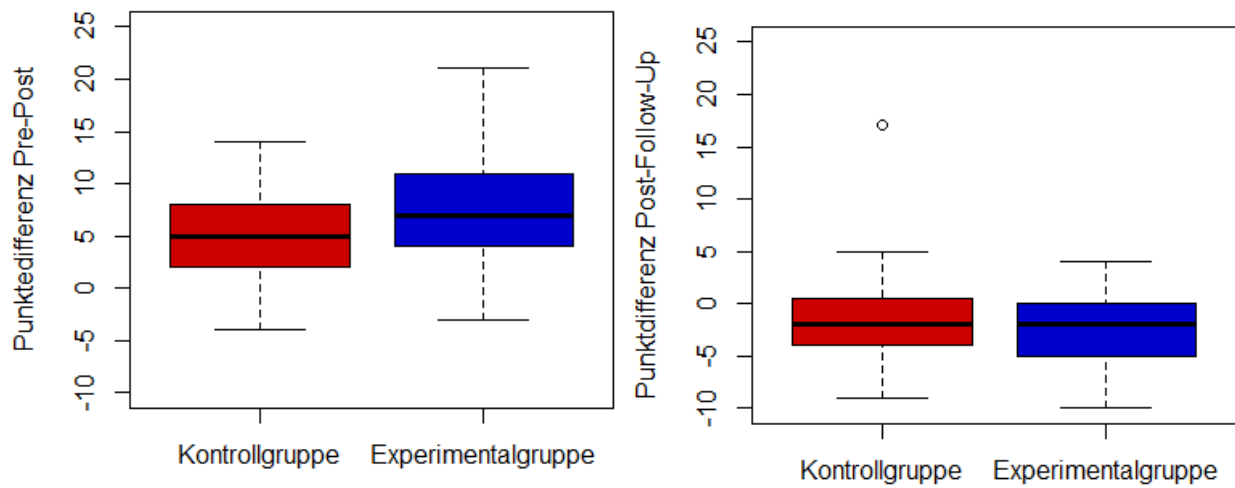


Abbildung 49: Lernzuwachs: Auswertung der Experimente 2, 3, 5 und 6 in einer Punktdifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll-und Experimentalgruppe.

Forschungsfrage 2.2: Welche Bedeutung hat die Qualität des Protokolls für die Lernwirksamkeit?

(H2) Die Lernwirksamkeit basierend auf einer experimentellen Tätigkeit korreliert mit der Qualität des Protokolls, welches dazu angefertigt wurde.

(H3) Schüler, die ein gutes Protokoll anfertigen, weisen eine höhere Lernwirksamkeit auf als Schüler, die ein schlechtes Protokoll erstellen.

Überprüfung auf Zusammenhänge zwischen Lernwirksamkeit und Protokollqualität

Zur Beantwortung dieser Forschungsfrage wurde exemplarisch ein Experiment mit dem dazugehörigen Protokoll ausgewählt. Die Begründung der Auswahl liegt in den Ergebnissen der Lernwirksamkeit. Der Lernzuwachs von Pre- zu Post-Test von Experiment 2 „Welches Gas entsteht beim Sprudeln der Brause?“ zeigt die größte Spanne der Lernwirksamkeit zwischen Minimum und Maximum und somit sowohl gute als auch schlechte Ergebnisse in der Lernwirksamkeit dieses Experimentes. Der Leistungstest besteht bei Experiment 2 aus sechs Wissens-Items und weist beim Lernzuwachs einen Maximalwert von 6 bis zu einem Minimalwert bei -3 der zu beantworteten Fragen auf. Der Mittelwert aller Probanden ist ein Lernzuwachs von $\chi = 1,62$ Fragen ($\sigma = 2,22$). Das Bewertungssystem für Protokolle wurde auf Experiment 2 adaptiert, um eine Bewertung dieser Protokolle vornehmen zu können (Kategoriensystem Experiment 2 im Anhang VII). Der Mittelwert der Protokollqualität aller Protokolle ($n = 51$) liegt bei $\chi = 1,58$ ($\sigma = 0,68$). Der Maximalwert in der Qualität ist 3, der Minimalwert bei 0. Die Kategorie Repräsentationsform wird deskriptiv verwendet und zeigt, dass insgesamt keine Tabellen, 2 Skizzen, 38 Texte und 11 Protokolle in Text und Skizze angefertigt wurden. Die Protokollqualität als auch die Ergebnisse zur Lernwirksamkeit sind bei diesem Experiment sehr heterogen.

Die Korrelation der Merkmale „Lernzuwachs Experiment 2“ und „Protokollqualität Experiment 2“ liegt bei $r = 0,08$ ($p\text{-Wert} = 0,45$). Dieser Wert liegt unter der angegebenen Definition für einen schwachen Zusammenhang. Auch weitere Merkmale des Protokolls, zum Beispiel einzelne Kategorien der Protokollqualität (Gliederung, Produkt, Beziehung), wurden auf einen Zusammenhang zur Lernwirksamkeit getestet. Die Korrelation von Gliederung ($r = 0,12$, $p = 0,246$), Produkt ($r = 0,14$, $p = 0,17$) und Beziehung ($r = 0,09$, $p = 0,35$) sind schwach positiv bis keinem linearen Zusammenhang zuzuordnen.

Einflüsse auf die Protokollqualität

Die Ergebnisse der Varianzanalyse in Tabelle 21 zeigen die Aufklärung der Unterschiede in der Protokollqualität von Experiment 2. Es gibt keine Unterschiede in der Qualität durch die Kontrollvariablen Migrationshintergrund, Nawi-Note, Deutschnote, Lernzuwachs und Motivation. Ein geringer Unterschied in der Protokollqualität kann durch das Geschlecht erklärt werden. Ein hoch signifikanter Unterschied (p -Wert < 0.01 **) der Protokollqualität kann durch die Zugehörigkeit der Probanden zur Kleingruppe erklärt werden.

Tabelle 21: Ergebnisse der Varianzanalyse der abhängigen Variable Protokollqualität mit unabhängigen Variablen. Die Fehlerwahrscheinlichkeit p (mit F-Wert, Freiheitsgrad) und die Effektgröße f nach COHEN sind angegeben.

Unabhängige Variable	p- und Effektgröße
Geschlecht	F(46)=3.14, p=0.08(.) f=0.26
Migrationshintergrund	F(46)=2.22 p=0.14 f=0.22
Nawi-Note	F(46)=0.42, p=0.79 f=0.12
Deutsch-Note	F(46)=0.88, p=0.49 f=0.29
Kleingruppe	F(46)=3.79, p<0.01 (**) f=1,28
Lernzuwachs Experiment 2	F(46)=0.04, p=0.97 f=0.13
Motivation	F(46)=1.55, p=0.88 f=0.02

Um das Ergebnis der Varianzanalyse der Kleingruppe besser zu verdeutlichen, sind in Abbildung 50 die Protokolle zweier Probanden einer Kleingruppe zu einem Experiment dargestellt. Die Protokolle sind in Gliederung und Inhalt identisch.

Zeichne und schreibe deine Ergebnisse so auf, dass du anderen Kindern das Experiment erklären kannst.



Es entsteht das Löschgass
Das blaue Wasser wird gelb

⇒ Kohlenstoffdioxid

Teste zuerst die Reaktion des Wassers indem sich 10 Tropfen Bromthymolblau befinden, auf das Hinzufügen der verschiedenen Gase: Kohlenstoffdioxid, Luft und Distickstoffoxid. So findest du heraus welches Gas bei der Brause entsteht.

Zeichne und schreibe deine Ergebnisse so auf, dass du anderen Kindern das Experiment erklären kannst.



- Es entsteht das Löschgass

⇒ Kohlenstoffdioxid

Teste zuerst die Reaktion des Wassers, indem sich 10 Tropfen Bromthymolblau befinden, auf das Hinzufügen der verschiedenen Gase: Kohlenstoffdioxid, Luft und Distickstoffoxid. So findest du heraus welches Gas bei der Brause entsteht.

Abbildung 50: Protokolle zu Experiment 2 der ID-Nummern 257207 und 254109. Die beiden Probanden arbeiteten in einer Kleingruppe.

3.3.7 Diskussion der Ergebnisse

Bedeutung der Ergebnisse für Forschungsfrage 2.1

„Welche Bedeutung hat das Protokollieren für die Lernwirksamkeit?“

Als Element des naturwissenschaftlichen Problemlösens wird dem Experimentieren eine entscheidende Rolle in Bezug auf die Lernwirksamkeit zugeschrieben (HOFSTEIN & LUNETTA, 2004). Mit dieser Forschungsfrage sollte untersucht werden, ob der Teilschritt des Protokollierens beim Experimentieren bereits einen Einfluss auf die Lernwirksamkeit hat. Dabei galt es durch einen Gruppenvergleich in den Leistungstests folgende Hypothese zu prüfen:

(H1) Das Anfertigen eines Protokolls während des Experimentierens erhöht die Lernwirksamkeit der Schüler im Vergleich zu Schülern, die kein Protokoll anfertigen.

Bezogen auf die Abschnitte „Allgemeine Lernwirksamkeit“ und „Vergleich der Lernwirksamkeit der Interventionsgruppen“ stellt der positive Wert im Mittelwertvergleich und im Lernzuwachs dar, dass etwas gelernt wurde. Ein negativer Wert in diesem Bereich veranschaulicht, dass Inhalte vergessen wurden. Die Mittelwerte und die Ergebnisse der t-Tests für den allgemeinen Lernerfolg der Experimente 1, 2, 3, 5 und 6 zeigen, dass sowohl von Pre- zu Post-Test als auch von Pre- zu Follow-Up-Test deutlich mehr Fachwissensfragen richtig beantwortet werden konnten. Zwar ist ein geringeres Fachwissen von Post- zu Follow-Up-Zeitpunkt zu verzeichnen, dies war aber zu erwarten. Das Fachwissen fällt nicht unter den Ausgangswert des Pre-Tests. Der geringe Lernerfolg bei Experiment 4 zeigt, dass dieses Experiment und die Fachinhalte zu anspruchsvoll für diese Altersstufe waren. Deshalb wurde dieses Experiment für weitere Vergleiche ausgeblendet. Wird der Lernerfolg über den Gesamtwert der Experimente verglichen, zeigt sich, dass Fachwissen durch die Lerneinheit aufgebaut wurde. Die entwickelte Unterrichtseinheit kann als lernförderlich für den Erwerb von Fachwissen bezeichnet werden und ist somit geeignet für einen Vergleich der Interventionsgruppen.

Durch die Aufklärung der Kontrollvariablen konnte gezeigt werden, dass bei den Experimenten 2, 3, 5 und 6 bei dem Lernerfolg keine Abhängigkeit durch weitere Variablen vorliegt. Nur bei Experiment 1 zeigt der Lernerfolg eine Abhängigkeit zu den Kontrollvariablen. Deshalb wurde auch dieses Experiment für weitere Vergleiche herausgenommen. Einen großen Effekt auf den Gesamtwert im Lernerfolg hat die Zugehörigkeit zur Kleingruppe. Die Unterschiede in den Kleingruppen sind signifikant in Bezug auf den Lernerfolg. Die Zusammenar-

beit in der Gruppe kann sowohl einen positiven Effekt als auch einen negativen Effekt auf den Lernerfolg der Schüler haben. Mit einer kleineren Gruppengröße oder Einzelarbeit könnte dieses Problem vermieden werden. Jedoch erhöht sich dadurch vor allem die Anzahl der Materialien.

Die Überprüfung von Fachinhalten einer selbst konzipierten Lerneinheit ist zur Beantwortung zahlreicher Forschungsfragen notwendig. Deshalb ist das Entwickeln von Leistungstest zu den Inhalten eine gängige Vorgehensweise. Die Experimente dieser Einheit orientieren sich an speziellen thematischen Inhalten und nicht an allgemeinen Basiskonzepten der Chemie. Deshalb war es nicht möglich auf bestehende Tests zurück zu greifen. Eine Überprüfung des Tests auf Schwierigkeit und eine Überarbeitung der Antwortmöglichkeiten wurde nach der Pilotierung zwar vorgenommen und die Reliabilität sowie Trennschärfe des Tests überprüft. Die unzureichende Trennschärfe schränkt die Interpretierbarkeit dieser Daten im Hinblick auf die Leistung der Schüler ein. Das geschlossene Aufgabenformat des Leistungstests bringt zwar eine höhere Auswertungsobjektivität mit sich, jedoch steigt durch die Multiple Choice Antworten auch die Ratewahrscheinlichkeit der Schüler. Es ist durchaus möglich, dass einige Schüler nicht die richtige Antwort wussten und diese aber geraten haben. Diese Problematik hätte umgangen werden können, indem die Schüler bei der Beantwortung der Fragen angeben, ob sie sich bei der Beantwortung der Fragen sicher sind (THEYSEN, 2014). Die Situation während der Erhebung der Daten muss an dieser Stelle auch erwähnt werden. Trotz größter organisatorischer Bemühungen seitens der Schulleitung, ist es während des Schulalltags kaum möglich eine vergleichbare Leistungsmessung durchzuführen. Die zeitliche Lage der Datenerhebung während des Schultages, vor oder nach wichtigen Ereignissen für die Schüler oder einfach nur das heiße Sommerwetter können die Stimmung und Motivation der Klasse schnell ändern. Ein Datenerhebungsprotokoll hält solche Ereignisse zwar fest und gestattet Rückschlüssen auf die Ergebnisse, ändert jedoch die Situation vor Ort nicht. Die Art und Weise der Datengenerierung muss bei der Interpretation der Ergebnisse immer berücksichtigt werden.

Die Ergebnisse des Fachtests beider Interventionsgruppen zu Beginn der Untersuchung zeigen, dass das Vorwissen bei allen Experimenten gleich ist. Es gelten gleiche Startvoraussetzungen für beide Gruppen. Kleine Unterschiede in der Lernwirksamkeit der beiden Interventionsgruppen sind sowohl zum Post- als auch zum Follow-Up-Zeitpunkt zu erkennen. Die Experimentalgruppe liegt in der Lernwirksamkeit bei allen Experimenten und auch im Ver-

gleich des Lernzuwachses immer etwas über der Kontrollgruppe. Die Varianzanalysen der Interventionsgruppen zeigen jedoch in den einzelnen Experimenten kein signifikantes Ergebnis. Es gibt somit keinen signifikanten Unterschied im Lernerfolg der beiden Gruppen, außer bei dem Pre-Post-Vergleich von Experiment 6. In diesem Experiment ist der Lernerfolg der Experimentalgruppe signifikant und mit einem mittleren Effekt größer als bei der Kontrollgruppe. Der Vergleich des Lernerfolges über alle sechs Experimente hinweg summiert die kleinen Unterschiede der einzelnen Experimente auf und zeigt dadurch ein deutlicheres Ergebnis: Der Lernerfolg der Experimentalgruppe ist über die drei Messzeitpunkte signifikant mit einem geringen Effekt größer als bei der Kontrollgruppe. Auch der Lernzuwachs von Pre- zu Post-Test ist bei der Experimentalgruppe signifikant mit einem mittleren Effekt größer als bei der Kontrollgruppe.

Die allgemeine positive Lernwirksamkeit der gesamten Stichprobe zeigt zunächst, dass auch in der Kontrollgruppe ohne das Schreiben von Protokollen eine Lernwirksamkeit zu verzeichnen ist. Das Anfertigen von Protokollen ist nicht zwingend notwendig, um etwas zu lernen. Auch die teilweise große Spanne in der Lernwirksamkeit innerhalb der Experimentalgruppe zeigt, dass das Anfertigen von Protokollen nicht bei jedem Schüler direkt zur Lernwirksamkeit führt. Über die gesamte Lerneinheit hinweg, ist die Intervention des Protokollierens allgemein als förderlich für die Lernwirksamkeit zu betrachten. Dieses Ergebnis scheint zunächst nicht überraschend, da sich die Schüler durch das Protokollieren intensiver mit der Thematik des Experimentes beschäftigen. Ein Zitat von LACHMAYER, NERDEL & PRECHTL (2007) fasst dieses Ergebnis passend zusammen: *„Es zeigt sich, dass das eigenständige Konstruieren einen förderlichen Einfluss auf das Lernen haben kann, aber nicht muss.“* Die Effektivität des Schreibens könnte mit der Funktion des Protokolls für den Schreiber zusammenhängen. Nach BEERENWINKEL ET AL. (2016) werden drei mögliche Funktionen des Schreibens beschrieben: (1) Konservierende Funktion als Gedächtnisentlastung, (2) dokumentierende Funktion zum Festhalten der Ergebnisse für anderen und (3) epistemische Funktion zum Erlangen von Erkenntnissen durch das Schreiben. Der Arbeitsauftrag des Forscherheftes impliziert eine dokumentierende Funktion des Erarbeitungsprotokolls. Jedoch kann nicht davon ausgegangen werden, dass die Schüler das Protokoll nicht auf eine andere Art nutzen. Wird das Schreiben des Protokolls zur Erkenntnisgewinnung genutzt, dann kann daraus eine höhere Lernwirksamkeit hervorgehen, als bei den anderen Funktionen des Schreibens.

Die Auswertung der Protokolle zeigt jedoch, dass auch falsche Ergebnisse in den Protokollen verschriftlicht wurden. Das Schreiben von Protokollen garantiert noch keine richtigen Ergebnisse oder Schlussfolgerungen aus dem Experiment. Zusätzlich könnte es in diesem Fall sogar zu einer kontraproduktiven Wirkung des Protokollierens kommen, indem sich falscher Rückschluss verfestigt. Dies könnte eine Erklärung für die manchmal ausbleibende Lernwirksamkeit innerhalb der Experimentalgruppe liefern. Die Besprechung der Ergebnisse mit der Gruppe würde dieses Phänomen ausschließen. Zu Beginn der Untersuchung gab es bereits die Überlegung, die Protokolle am Ende der Stunde mit den Schülern gemeinsam zu kontrollieren, um richtige Ergebnisse zu gewährleisten. Diese Intervention wurde jedoch nicht durchgeführt, um den Vergleich zwischen den Gruppen fair zu gestalten. In der Kontrollgruppe gab es im Ablauf der Stunde keine Möglichkeit falsche Ergebnisse zu korrigieren. Außerdem wäre bei einer solchen Maßnahme die Besprechung der Ergebnisse in den Fokus der Intervention gerückt und nicht das Protokollieren. Der letzte Schritt des Erkenntnisprozesses, das Ziehen einer Schlussfolgerung aus den Ergebnissen, wurde deshalb nicht durch das Design gewährleistet, sondern musste durch Eigenleistung der Schüler erfolgen. Eine weitere Möglichkeit die Spanne im Lernerfolg zu erklären, bietet der Zeitpunkt an dem das Protokoll anfertigt wird. Das Protokollieren findet nicht nur in der Protokollphase am Ende der Stunde statt, sondern auch während des Experimentierens. Die Kleingruppe teilt sich die Aufgaben: Ein Teil experimentiert, der andere Teil protokolliert die Ergebnisse. Das Schreiben während des Experimentes teilt die Aufmerksamkeit der Schüler. Es kann sich entweder vollständig auf das Experiment oder auf das Protokoll konzentriert werden.

Bei der Interpretation der Ergebnisse gilt zu beachten, dass die Intervention nur fünf Minuten pro Experiment einnahm. Bei der Gesamtdauer von sechs Experimenten, entsprach die Intervention 10% der Gesamtzeit von 240 Minuten. Für eine derart geringe Interventionsmaßnahme zeigt sich dennoch in Summe über alle Experimente ein signifikanter Unterschied zur Kontrollgruppe mit einem geringen bis mittleren Effekt. Durch eine größere Stichprobe könnten ebenso deutlichere Ergebnisse der Gruppenunterschiede erzielt werden. Die Berechnung der stichprobenunabhängigen Effektstärke zeigt bereits, wie deutlich die Ergebnisse sein können.

Bedeutung der Ergebnisse für Forschungsfrage 2.2

„Welche Bedeutung hat die Qualität des Protokolls für die Lernwirksamkeit?“

Um die Forschungsfrage bezüglich des Zusammenhangs von Lernwirksamkeit und Qualität des Protokolls zu beantworten, wurde ein Zusammenhangsmaß der beiden Werte berechnet. Es galt folgende Hypothesen zu überprüfen:

(H2) Die Lernwirksamkeit eines Experimentes steht in Verbindung mit der Qualität des Protokolls, welches dazu angefertigt wurde.

(H3) Schüler, die ein gutes Protokoll anfertigen, haben eine höhere Lernwirksamkeit als Schüler, die ein schlechtes Protokoll anfertigen.

Die Korrelation der Gesamtqualität der Protokolle mit der Lernwirksamkeit zeigt keinen linearen Zusammenhang der beiden Merkmale. Auch die Überprüfung auf Zusammenhang einzelner Kategorien der Protokollqualität und Lernwirksamkeit zeigen einen geringen positiven bis keinen linearen Zusammenhang. Beide Hypothesen werden dadurch falsifiziert.

Die Varianzanalyse ergab einen hochsignifikanten Unterschied in der Protokollqualität in den Kleingruppen. Dies bedeutet, dass die Arbeit in der Kleingruppe die Protokollqualität beeinflusst. Die Ergebnisse der Varianzanalyse unterstützen die Beobachtungen der Betreuer während der Intervention und während der Protokollauswertung. Wie Abbildung 51 zeigt, führte das Zusammenarbeiten der Schüler in den Kleingruppen beim Erstellen der Protokolle in den Protokollphasen zu oft identischen Protokollen. Die erstellten Protokolle zeigen deshalb nicht immer individuelle Leistungen auf und lassen sich deshalb auch nur schwer auf die individuelle Lernerfolgsleistung beziehen. In den Kleingruppen kommt es zur Absprache beim Schreiben der Protokolle, deshalb können die Protokolle nicht als individuelle Leistungen betrachtet werden. Auch die Arbeitsverteilung während des Experimentierens trägt zu diesem Phänomen bei. Innerhalb der Kleingruppe werden die Arbeitsweisen, wie Experimentieren und Protokollieren verteilt und anschließend die Ergebnisse ausgetauscht. Abbildung 51 zeigt dieses



Abbildung 51: Kleingruppe beim Experiment. Ein Schüler experimentiert, zwei Schülerinnen protokollieren.

häufig beobachtete Phänomen der Aufteilung der Methoden. Es kann deshalb nur ein eingeschränkter Rückschluss auf die individuelle Leistung der Protokolle getroffen werden. Zwar wurde eine Einzelarbeit beim Verschriftlichen gefordert, dies konnte jedoch nicht fortlau-

fend überprüft werden. Die Protokolle werden manchmal in Gruppenarbeit angefertigt, manchmal auch individuell. Eine Minimierung der Kleingruppen auf zwei Personen oder eventuell Einzelarbeit würden dieses Problem zwar beheben, wäre aber zum einen ein größerer Materialaufwand und zum anderen untypisch bei Schülerexperimentierphasen.

Es bleibt die Frage, wie Unterschiede in der Protokollqualität entstehen. Die Varianzanalyse zeigte, dass die Varianz der Protokollqualität durch die Aspekte Lernleistung, Note, Geschlecht und Motivation nicht erklärt werden kann. Weitere zu berücksichtigenden Faktoren könnten die kognitive Fähigkeit und eine Schreibmotivation während der Protokollphase sein. Allgemein scheint die Methode zur Bewertung von Protokollen in sich stabil. Jedoch ist schwierig diese Fertigkeit einem bestehenden Konstrukt zuzuordnen. Die Einflussfaktoren auf das Protokollieren wurden durch die Ergebnisse von Teilstudie 2 minimiert, das Konstrukt und mögliche weitere Einflüsse sind jedoch noch nicht ausreichend untersucht. Deshalb ist es schwierig weitere Kontrollvariablen mit in das Design einzubeziehen. Fortfolgende Studien müssen zeigen, ob das Protokoll nur als Erinnerungstütze dient oder ein Abbild der Vorgehensweise ist.

4 Zusammenfassung und Ausblick

In der Schule ist die Tätigkeit des Protokollierens durch die Bildungsstandards im naturwissenschaftlichen Unterricht verankert. Die Kompetenzbereiche Erkenntnisgewinnung und Kommunikation weisen beide den Anspruch auf, ein Protokoll anzufertigen. Im Bereich Erkenntnisgewinnung ergibt sich durch das experimentelle Arbeiten stets die Möglichkeit Daten, in der Form eines Protokolls, zu erheben und auszuwerten. Die Teilbereiche und Aspekte des Kompetenzbereichs Kommunikation können durch die Tätigkeit des Protokollierens bedient werden. Die allgemeine Bedeutung des Schreibens im naturwissenschaftlichen Unterricht zeigt eine weitere Bedeutung des Protokollierens auf. Obwohl der Schreibvorgang eine unbeliebte Tätigkeit ist, werden dadurch Lern- und Denkprozesse angestoßen. Es gibt zahlreiche Möglichkeiten Protokolle im Unterricht einzusetzen. Zu Beginn der Arbeit wurden sechs Protokollformen aufgezeigt und deren Einsatzgebiete erläutert. Je nach Protokollform gibt es unterschiedliche Hilfestellungen zum Erlernen der Protokolliertechniken. Diese Hilfestellungen sollen den Schülern die unbeliebte Tätigkeit des Protokollierens erleichtern.

Das Lernen in Schülerlaboren hat in den letzten zehn Jahren als Ergänzung zum Unterricht und den naturwissenschaftlichen Arbeitsweisen an Bedeutung gewonnen. Das Projekt PROLAB–Protokollieren im Labor untersuchte im Rahmen von drei Teilstudien die Bedeutung des Protokolls für den naturwissenschaftlichen Erkenntnisweg im außerschulischen Bereich. In Teilstudie 1 wurde zunächst der Status Quo des Umgangs mit dem Protokollieren an außerschulischen Lernorten erhoben. Die Einflussfaktoren auf das Produkt des Protokolls und bei der Tätigkeit des Protokollierens standen in Teilstudie 2 im Fokus. In einer dritten Teilstudie wurden die Bedeutung des Protokollierens für die Lernwirksamkeit und der Zusammenhang der Qualität der Protokolle mit der Lernwirksamkeit untersucht. Die Studien wurden ausschließlich im Schülerlabor und nicht in der Schule durchgeführt. Die Ergebnisse und die daraus resultierenden Empfehlungen sind deshalb zunächst nur für Schülerlabore geltend zu machen. Durch die Methodenkritik werden weitere Einschränkungen der Ergebnisse festgestellt. Es lässt sich ebenso keine Verallgemeinerung auf weitere Klassenstufen oder Fächer veranschlagen. Die Interpretation ist für eine sechste Klassenstufe im Fach Naturwissenschaften in Rheinland-Pfalz gültig. Um allgemeingültige Aussagen treffen zu können, müssten weitere Untersuchungen stattfinden.

Teilstudie 1 wurde mit dem Bundesverband der Schülerlabore (LELA E.V.) durchgeführt. In der Online-Fragebogenstudie erfolgte eine Rückmeldung von N=87 Schülerlaboren. Die Ergebnisse veranschaulichen, dass die Häufigkeit des Protokollierens vom Alter der teilnehmenden Schüler abhängt. In älteren Jahrgangsstufen wird das Protokoll vielmehr als Werkzeug beim Experimentieren eingesetzt als bei jüngeren Schülern. Beispielsweise werden in Schülerlaboren mit der Zielgruppe der Orientierungsstufe nur in 18% Protokolle optimal eingesetzt. Das Protokollieren im Schülerlabor ist jedoch zusätzlich von der Einstellung der Schülerlaborleitung abhängig. Sobald das Protokollieren vom jeweiligen Schülerlabor als notwendig erachtet wird, wird verstärkt zum Protokollieren aufgefordert und Materialien zum Protokollieren eingesetzt. Diese Studie zeigt, dass das Protokollieren nicht nur im schulischen Bereich sondern auch in Schülerlaboren als Methode und Arbeitsweise der Erkenntnisgewinnung eingesetzt wird. Die Übersicht der unterschiedlichen Definitionen und Anwendungsbereiche veranschaulicht die Bedeutung des Protokolls im Kontext Schule. Die Hilfestellungen zum Erlernen der Protokolltechniken zeigen jedoch auch, dass das Protokollschreiben in der Schule sich als schwierig herausstellt.

Teilstudie 2 zeichnet sich durch eine explorative Herangehensweise aus. Hier konnten wertvolle Einsichten in dieses zuvor unbekannte Thema gewonnen werden. Eine Vergleichsstudie in einem interdisziplinären Forscherverbund untersuchte die Tätigkeit des Protokollierens von N=180 Schülern der sechsten Klasse. Durch die Auswertung von Protokollen, Videoaufnahmen und Interviews konnten fünf Einflussfaktoren identifiziert werden, die sich auf das Protokollieren von Schülern auswirken. Die Merkmale Experiment, Lerngegebenheiten, Vorkenntnisse, Gruppendynamik und Einstellung beeinflussen die Darstellungsweise der Protokolle in verschiedenen Ausprägungen. Diese Teilstudie ermöglichte einen Einblick in unterschiedliche Protokollierweisen, Gruppenprozesse beim Protokollieren und mögliche Bewertungsmethoden von Protokollen. Ebenso konnten Schwierigkeiten beim Schreiben von Protokollen (zum Beispiel Heftgestaltung, Aufgabenstellen und weiteres) identifiziert werden. Durch die Erkenntnisse und kritische Reflexion wurde eine gezielte Planung für Teilstudie 3 möglich. Ebenso konnte durch die Zusammenarbeit mit der Fachdidaktik Mathematik ein neuer Protokollbegriff definiert werden. Erarbeitungsprotokolle grenzen sich von den übrigen Begrifflichkeiten ab, indem nur die Arbeitsprozesse und Ergebnisse während des Experimentierens festgehalten werden.

Teilstudie 3 untersuchte die Bedeutung des Protokollierens für die Lernwirksamkeit, um die Notwendigkeit nochmal explizit hervorzuheben. In einer Vergleichsstudie mit N=115 Schülern wurde die Lernwirksamkeit von Schülern ohne Protokolliertätigkeit und Schülern mit Protokolliertätigkeit gegenübergestellt. Die Schüler dieser Studie zeigen allgemein eine Lernwirksamkeit durch die Lerneinheit auf. Auch ohne den Teilschritt des Protokollierens tritt ein Lernerfolg auf. Durch das Protokollieren wird jedoch geringfügig (p -Wert=0.016 *, $\delta=0.12$) über drei Messzeitpunkte mehr gelernt. Daraus kann man schließen, dass das Protokollieren die Konstruktion von Wissen fördert. Der Einfluss der Zugehörigkeit zu einer Kleingruppe ($p=0.0017$ **, $\eta^2 = 0.23$ und $p < 0.01$ **, $f=0.95$) auf den Lernerfolg ist jedoch ebenso zu verzeichnen. In dieser Studie konnte zugleich eine Bewertungsmethode für Erarbeitungsprotokolle etabliert werden (vgl. ENGL ET AL., 2015). Die Qualität der Protokolle wird mit Hilfe von vier Kategorien ermittelt. Die Kategorien können an den Inhalt angepasst werden und sind somit auf weitere Protokolle übertragbar. Die Zusammenhangsprüfung der Lernwirksamkeit mit der Qualität der Protokolle ergab keine Korrelation. Dieses Ergebnis konnte hauptsächlich durch das gemeinsame Protokollieren in der Gruppe und fehlerhaftes Protokollieren erklärt werden. Das Protokoll wurde als ein notwendiger Bestandteil der Bildungsstandards und der naturwissenschaftlichen Erkenntnisgewinnung theoretisch als auch empirisch darstellt.

Durch die Ergebnisse und Interpretation von PROLAB können Empfehlungen für Schule abgeleitet werden, welche jedoch nicht empirisch geprüft wurden. Beachtet man die Einflussfaktoren des Protokollierens, kann der Vorgang für den Protokollanten angenehmer gestaltet werden. Zum Beispiel durch die Gestaltung des Protokollierbereichs nahe der Experimentvorschrift oder dem Einräumen von ausreichend Zeit für das Protokollieren. Ebenso lohnt es sich das Protokollieren trotz des negativen Rufes nicht zu vernachlässigen. Eine Nachbesprechung des Protokolls wird als notwendig erachtet, um keine falschen Schlussfolgerungen zu gewähren, die sich bei den Schülern verfestigen könnte. Gegen alle Erwartungen scheint die Qualität des Protokolls nicht entscheidet für die Lernwirksamkeit zu sein. Solange nicht geklärt ist, was hinter den wiedergegebenen Inhalten eines Protokolls steckt, ist das eigenständige Erstellen der Protokolle ohne Hilfsmittel und Vorgaben am gewinnbringendsten.

Durch die Durchführung weiterer Studien könnten die Ergebnisse dieser und weiterer Studien gezielt in die Schule übertragen werden und somit Schülern und Lehrern das Protokollieren im Unterricht erleichtern. Mit weiteren Arbeiten könnten zusätzliche Aspekte des Pro-

protokollierens untersucht werden. Vor allem durch die Einteilung der Schüler in Kleingruppen, war es nahezu unmöglich die individuelle Leistung der Schüler im Hinblick auf das Erstellen eines Protokolls zu erörtern. Jedoch wären genau diese individuellen Leistungen ohne Gruppeneinfluss interessant für eine weitere Darstellung des Protokollierens. Ebenso könnte der Frage nachgegangen werden, was tatsächlich von den Schülern protokolliert wird. In zusätzlichen Interviews mit den Protokollanten wurden mögliche Hinweise darauf gesammelt, dass das Protokoll als unterschiedliche Lernstrategie genutzt wurde. Für einige Schüler diente das Protokoll als Erinnerungstütze von Inhalten, die ihnen vorher unbekannt waren und sie ergänzten diese in den Interviews mit ihrem Wissen. Andere Schüler versuchten nach Möglichkeit alle gegebenen Inhalte zu protokollieren. In einer weiteren Studie könnten diese individuellen Strategien erörtert und klassifiziert werden. Ein solches Ergebnis hätte Auswirkungen auf die Idee Protokolle nach ihrer Qualität zu beurteilen. Eine weitere Möglichkeit die Strategien hinter dem Protokollieren zu erklären, wäre die Entwicklung einer spezifischen Fertigkeit. Gibt es ein Konstrukt oder vielleicht eine Kompetenz die aussagt, wie gut ein Schüler Protokolle schreiben kann? In diesem Fall sollten Längsschnittstudien angedacht werden, die diese Protokollierfähigkeit der Schüler über verschiedene Jahrgangsstufen verfolgen. Dadurch wäre es möglich Aussagen über die Entwicklung und Förderung dieser Kompetenz zu machen. Ebenso wären Studien des Protokollierens mit einem Fokus im naturwissenschaftlichen Schulunterricht empfehlenswert.

Zusammenfassend lässt sich eine nachweisliche Bedeutung des Protokollierens für den naturwissenschaftlichen Erkenntnisweg empirisch belegen. Um das Schreiben von Protokollen zu erleichtern, müssen die Darstellungsart und die Einflussfaktoren beachtet werden. Die Notwendigkeit in Bezug auf die Lernwirksamkeit ist ganzheitlich zu betrachten. Denn nur das Zusammenspiel einzelner Teilprozesse der Erkenntnisgewinnung führt zur Lernwirksamkeit. Um den Gesamtprozess für Lernende möglichst effektiv zu gestalten, muss die Bedeutung der einzelnen Teilschritte des Erkenntnisprozess dargestellt gemacht werden.

Literatur

BAKEMAN, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37 (3), 379-384.

BÄURLE, W., FRÖCHTENICHT, E., GANZ, G., GIETZ, P., HEITLAND, W., HOPPE, B., MÜLLER, O., PEPPMEIER, R., SCHLEUSENER, P., WÄCHTER, M., WEIZEL, B., WILLMER-KLUMPP, C. & WOLF, U. (2006). *Prisma Chemie 7-10*. Stuttgart: Ernst Klett Verlag.

BECKER-MROTZEK, M. & BÖTTCHER, I. (2011). *Schreibkompetenz entwickeln und beurteilen, Praxishandbuch für die Sekundarstufe I und II*. Berlin: Cornelsen Verlag.

BECKER-MROTZEK, M., SCHRAMM, K., THÜRMAN, E. & VOLLMER, H.J. (2013). Sprache im Fach: Einleitung. In Becker-Mrotzek, M., Schramm, K., Thürmann, E. & Vollmer, H.J. (Hrsg.), *Sprache im Fach, Band 3* (S. 7-14). Münster: Waxmann Verlag.

BEERENWINKEL, A., HEFTI, C., LINDAUER, T., & SCHMELLENTIN, C. (2016). Schreiben im Chemieunterricht. *CHEMKON*, 23(1), 19-24.

BERGE, O. E. & VOLKMER, M. (2002). Schülerexperimente als Testsituation – Hinweise zur Konzeption und Bewertung. *Naturwissenschaften im Unterricht Physik*, 71-72, 10-11.

BERGELER, E. (2009). *Lernen durch eigenständiges Schreiben von sachbezogenen Texten im Physikunterricht*. Verfügbar unter www.qucosa.de/fileadmin/data/qucosa/documents/2278/Doktorarbeit.pdf letzter Zugriff [27.06.2015].

BERGELER, E. & POSPIECH, G. (2008). Schreiben im Physikunterricht. In Höttecke, D. (Hrsg.), *Kompetenzen, Kompetenzmodelle, Kompetenzentwicklung, Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Essen 2007* (S. 320-32). Münster: Lit.

BERTHOLD, K., NÜCKLES, M., & RENKL, A. (2004). Writing learning protocols: Prompts foster cognitive and metacognitive activities as well as learning outcomes. In Gerjets, P., Elen, J. Joiner, R. & Kirschner, P. (Hrsg.), *Instructional design for effective and enjoyable computer-supported learning* (S. 193–200). Tübingen: Knowledge Media Research Center.

BERTHOLD, K., NÜCKLES, M. & RENKL, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17(5), 564-577.

BIBLIOGRAPHISCHES INSTITUT GMBH (2013). *Duden (online Version)*. Verfügbar unter <http://www.duden.de/rechtschreibung/Protokoll> letzter Zugriff [28.5.2015].

- BÖTTCHER, I. & BECKER-MROTZEK, M. (2009). *Texte bearbeiten, bewerten und benoten*. Berlin: Cornelsen Verlag.
- BORTZ, J. & DÖRING, N. (2009). *Forschungsmethoden und Evaluation*. Heidelberg: Springer Verlag.
- BORTZ, J. & SCHUSTER, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer Verlag.
- BRITTON, J. (1970). *Language and learning*. London: Allen Lane.
- BÜCKMANN, M. & DUIT, R. (2014). Videobasierte Analyse unterrichtlicher Sachstrukturen. In Krüger, D. Schecker, H. & Parchmann, I. (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S.189-194). Heidelberg: Springer Spektrum.
- BRUENING, H.G. (1990). Das Versuchsprotokoll. *Physica didactica*, 17, 101-109.
- BRUNER, J.S. (1971). Über kognitive Entwicklung. In J.S. Bruner, R.R. Olver & P.M. Greenfield (Hrsg.), *Studien zur kognitiven Entwicklung: Eine kooperative Untersuchung am Center for Cognitive Studies der Harvard-Universität* (S. 21-96). Stuttgart: Klett Verlag.
- BUSCH, H. & RALLE, B. (2013). Diagnostik und Förderung fachsprachlicher Kompetenzen im Chemieunterricht. In Becker-Mrotzek, M., Schramm, K., Thürmann, E., Vollmer, H.J. (Hrsg.), *Sprache im Fach, Band 3* (S.277-294). Münster: Waxmann Verlag.
- COX, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, 9, 343-363.
- DÖRFER, W. (2003). Protokolle und Diagramme als ein Weg zum diskreten Funktionsbegriff. In Hoffmann, M.H.G. (Hrsg.), *Mathematik verstehen: Semiotische Perspektiven* (S. 78–94). Hildesheim: Franzbecker Verlag.
- DÖRFLER, W. (2012). Handlungen und Protokolle, Abstraktion und Verallgemeinerung. In Sonntag, M., Hörmannseder, R. (Hrsg.), *Informatik: Von Anfang an. Festschrift für Jörg R. Mühlbacher* (S. 45-52). Linz: Tauber Verlag.
- DURST, R.K., NEWELL, G.E. (1989). The Uses of Function: James Britton's Category System and Research on Writing. *Review of Educational Research*, 59, 375–394.
- EID, M., GOLLWITZER, M. & SCHMITT, M. (2011). *Statistik und Forschungsmethoden*. Weinheim: Beltz-Verlag.
- EMDEN, M. & SUMFLETH, E. (2012). Prozessorientierte Leistungsbewertung des experimentellen Arbeitens. Zur Eignung einer Protokollmethode zur Bewertung von Experimentierprozessen. *Der mathematische und naturwissenschaftliche Unterricht*, 65(2), 68-75.

- ENGL, L., SCHUMACHER, S., SITTER, K., GRÖßLER, M., NIEHAUS, E., RASCH, R., ROTH, J. & RISCH, B. (2015). Entwicklung eines Messinstrumentes zur Erfassung der Protokollierfähigkeit—initiiert durch Video-Items. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 223-229.
- FISCHER, H. E., KLEMM, K., LEUTNER, D., SUMFLETH, E., TIEMANN, R. & WIRTH, J. (2003). Naturwissenschaftsdidaktische Lehr-Lernforschung: Defizite und Desiderata. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 179-209.
- FLINTJER, B. (1998). Der handlungsinstruierende symbolorientierte Chemieunterricht. *CHEM-KON*, 5 (2), 108-109.
- FELLOWS, N. (1994). A window into thinking: Using student writing to understand conceptual change in science learning. *Journal of Research in Science Teaching*. 31(9), 985-1001.
- FORSTER, B., LANGER, S., VETRVSKY, C., & ANTON, M.A. (2011). Die Effizienz außerschulischer Lernorte. *Praxis der Naturwissenschaften - Physik in der Schule*, 60, 46-48.
- GALLIN, P. & RUF, U. (1991). *Sprache und Mathematik in der Schule. Auf eigenen Wegen zur Fachkompetenz*. Zürich: Verlag Lehrerinnen und Lehrer Schweiz (LCH).
- GALLIN, P. & RUF, U. (1993). Sprache und Mathematik in der Schule. Ein Bericht aus der Praxis. *Journal für Mathematik-Didaktik*, 14(1), 3-33.
- GALLIN, P. & RUF, U. (1995). Singuläre Schülertexte als Basis eines allgemeinbildenden Mathematikunterrichts. In Biehler, R., Heymann, H.W. & Winkelmann, B. (Hrsg.), *Mathematik allgemeinbildend unterrichten: Impulse für Lehrerbildung und Schule* (S. 58-82). Köln: Aulis.
- GLYNN, S.M. & MUTH, D.K. (1994). Reading and writing to learn science: Achieving scientific literacy. *Journal of Research in Science Teaching*, 31(9), 1057-1073.
- HELMKE, A. & HOSENFELD, I. (2004). Vergleichsarbeiten, Standards, Kompetenzstufen: Begriffliche Klärung und Perspektiven. In Wosnitza, M., Frey, A. & Jäger, R.S. (Hrsg.), *Lernprozess, Lernumgebung und Lerndiagnostik: Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert*, (S. 56-75). Landau: Verlag Empirische Pädagogik.
- HOFSTEIN, A. & LUNETTA, V.N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88(1), 28-54.
- HOLLIDAY, G., YORE, L. D. & ALVERMANN, D. E. (1994). The reading-science learning-writing connection: Breakthroughs, barriers and promises. *Journal of Research in Science Teaching*, 31(9), 877-893.
- HOPPE, P., KRÄMER, S. & REH, J. (2013). Wie keimt Kresse am besten - Durch schrittweises Vorgehen das Schreiben von Protokollen einführen. *Biologie 5-10*, 4, 12-15.

HÜBNER, S., NÜCKLES, M., & RENKL, A. (2007). Lerntagebücher als Medium des selbstgesteuerten Lernens – Wie viel instruktionale Unterstützung ist sinnvoll. *Empirische Pädagogik*, 21(2), 119-137.

HUG, B. & McNEILL, K.L. (2008). Use of First-hand and Second-hand Data in Science: Does data type influence classroom conversations? *International Journal of Science Education*, 30 (13), 1725–1751.

KEYS, C.W. (1994). The development of scientific reasoning skills in conjunction with collaborative writing assignments: An interpretive study of six nine-grade students. *Journal of Research in Science Teaching*, 31(9), 1003-1022.

KEYS, C.W., HAND, B., PRAIN, V. & COLLINS, S (1999). Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science. *Journal of Research in Science Teaching*, 36(10), 1065- 1084.

KLAHR, D. (2000). *Exploring Science: The Cognition and Development of Discovery Processes*. Massachusetts Institute of technology.

KLAHR, D. & DUNBAR, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12(1), 1-48.

KLIEME, E., AVENARIUS, H., BLUM, W., DÖBRICH, P., GRUBER, H. & PRENZEL, M. (2003). *Expertise zur Entwicklung nationaler Bildungsstandards*. Berlin: Bundesministerium für Bildung und Forschung.

KLOS, S., HENKE, C., KIEREN, C., WALPUSKI, M., & SUMFLETH, E. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen-zwei verschiedene Kompetenzen. *Zeitschrift für Pädagogik*, 54(3), 304-321.

KLUGE, F. (2011). *Etymologisches Wörterbuch der deutschen Sprache*. Berlin: De Gruyter.

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

KMK (2005b). *Beschlüsse der Kultusministerkonferenz – Bildungsstandard im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München, Neuwied: Luchterhand.

KMK (2005c). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München, Neuwied: Luchterhand.

KOBOW, I. (2015). *Entwicklung und Validierung eines Testinstrumentes zur Erfassung der Kommunikationskompetenz im Fach Chemie*. Berlin: Logos Verlag.

KOZMA, R. & RUSSELL, J. (2005a). Multimedia learning of Chemistry. In Mayer, R. (Hrsg.), *The Cambridge Handbook of Multimedia Learning*, (S. 409-428). Cambridge: Cambridge University Press.

KOZMA, R. & RUSSELL, J. (2005b). Students becoming chemists: Developing representational competence. In Gilbert, J. (Hrsg.), *Visualization in science education*, (S. 121-146). Dordrecht: Springer Netherlands.

KRAUS, M. E. & STEHLIK, S. (2008). Protokolle schreiben: Anregungen zur Auseinandersetzung mit einer problematischen Textsorte. *Naturwissenschaften im Unterricht – Physik in der Schule*, 104, 17-23.

KRÄMER, S. (2011). Brausepulver – eine prickelnde Mischung. Ein systematisches Training zum Protokolle schreiben. *Naturwissenschaften im Unterricht – Chemie*, 126, 23-28.

KUNTER, M., SCHÜMER, G., ARTELT, C., BAUMERT, J., KLIEME, E., NEUBRAND, M., PRENZEL, M., SCHIEFELE, U., SCHNEIDER, W., STANAT, P., TILLMANN, K.J. & WEIB, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.

LACHMAYER, S., NERDEL, C., & PRECHTL, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 161-180.

LANGER, I., SCHULZ VON THUN, F. & TAUSCH, R. (1981). *Sich verständlich ausdrücken*. München: Ernst Reinhardt Verlag.

LEDERMAN, N.G. (2007). Nature of Science: Past, Present and Future. In Abell, S.K & Lederman, N.G (Hrsg.), *Handbook of research on science education*, (S. 831–879). Mahwah, New Jersey: Erlbaum.

LIENERT, G. & RAATZ, U. (1998): *Testaufbau und Testanalyse*. Weinheim: Beltz-Verlag.

LÜCK, G. (2008). Naturphänomene sprachlich erfassen. *Naturwissenschaften im Unterricht - Chemie*, 106/107, 84-87.

LUHMANN, M. (2011). *R für Einsteiger – Einführung in die Statistiksoftware für die Sozialwissenschaften*. Weinheim: Beltz-Verlag.

MARSCHNER, J., THILLMANN, H., WIRTH, J., & LEUTNER, D. (2012). Wie lässt sich die Experimentierstrategie-Nutzung fördern?. *Zeitschrift für Erziehungswissenschaft*, 15(1), 77-93.

MAYER, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In Krüger, D. & Vogt, H. (Hrsg.), *Handbuch der Theorien in der biomedizinischen Forschung – Ein Handbuch für Lehramtsstudenten und Doktoranden*, (S. 178–186). Berlin/Heidelberg: Springer-Verlag.

- MAYRING, P. (2002). *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. Weinheim: Beltz Verlag.
- MAYRING, P. (2008). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim: Beltz-Verlag.
- MINISTERIUM FÜR BILDUNG, WISSENSCHAFT, JUGEND UND KULTUR (2010). *Rahmenlehrplan Naturwissenschaften für die weiterführenden Schulen in Rheinland-Pfalz*. Mainz.
- MOLL, M. (2003). Protokollieren heißt auch schreiben lernen. *Der Deutschunterricht*, 55(3), 71-80.
- MOOSBRUGGER, H., & KELAVA, A. (2008). *Testtheorie und Fragebogenkonstruktion*. Springer Berlin Heidelberg.
- NAWRATH, D., MAISEYENKA, V., & SCHECKER, H. (2011). Experimentelle Kompetenz - Ein Modell für die Unterrichtspraxis. *Praxis der Naturwissenschaften–Physik in der Schule*, 60(6), 42-49.
- NEUBERT, B. (2011). Kinder veranschaulichen ihre Denkprozesse und Lösungswege. Zur Entwicklung der allgemeinen Kompetenz des Darstellens im Mathematikunterricht. *Grundschulunterricht Mathematik*, 2, 4-6.
- NIESWANDT, M. (1998). Lernen im Chemieunterricht durch eigenständiges Schreiben. Fallanalysen, *Zeitschrift für Didaktik der Naturwissenschaften*, 2, 21-40.
- NÜCKLES, M., SCHWONKE, R., BERTHOLD, K., & RENKL, A. (2004). The use of public learning diaries in blended learning. *Journal of Educational Media*, 29(1), 49-66.
- PINTRICH, P. R., SMITH, D., GARCIA, T., & MCKEACHIE, W. (1991). *The motivated strategies for learning questionnaire (MSQL)*. Ann Arbor: University of Michigan.
- PFEILER, S., PRIEMER, B., UPMEIER ZU BELZEN, A. (2015). Der kritische Umgang mit Primär- und Sekundärdaten im naturwissenschaftlichen Unterricht. In Bernholt, S. (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht*, Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Bremen 2014 (S. 672-674). Kiel: IPN.
- PRAIN V. & HAND, B. (1996). Writing for learning in secondary science: Rethinking practises. *Teaching and Teacher Education*, 12 (6), 609–629.
- PRAIN, V. & TYTLER, R. (2013). Representing and learning in science. In R. Tytler, V. Prain, P. Hubber and B. Waldrip (Hrsg.), *Constructing Representations to Learn in Science* (S.1–14). Dordrecht: SensePublishers.
- PRECHTL, M. (2008). Versuchsprotokolle...einmal anders! Welche Chancen bieten Bildergeschichten?. In Höttecke, D. (Hrsg.), *Kompetenzen, Kompetenzmodelle, Kompetenzentwicklung*

lung, *Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Essen 2007* (S. 272-274). Münster: Lit.

PRENZEL, M. & PARCHMANN, I. (2003). Kompetenz entwickeln. Vom naturwissenschaftlichen Arbeiten zum naturwissenschaftlichen Denken. *Naturwissenschaften im Unterricht: Chemie*, 14, 15-19.

POSPECHILL, M. (2010). *Testtheorie, Testkonstruktion, Testevaluation*. Stuttgart: Reinhardt-Verlag.

RALLE, B. (2015). Sprachliche Heterogenität und fachdidaktische Forschung. In Bernholt, S. (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht, Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bremen 2014* (S. 4-18). Kiel: IPN.

REHM, M. (2006). Allgemeine naturwissenschaftliche Bildung – Entwicklung eines vom Begriff Verstehen ausgehenden Kompetenzmodells. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 22-44.

REHM, M. (2007). Das empirische Forschungsinstrument „Phänomenprotokoll“. In Höttecke, D. (Hrsg.), *Naturwissenschaftlicher Unterricht im internationalen Vergleich, Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bern 2006* (S. 269-271).

RHEINBERG, F.; VOLLMEYER, R. & BURNS, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47, 57–66.

RIESE, J. & REINHOLD, P. (2014). Entwicklung eines Leistungstests für fachdidaktisches Wissen. In Krüger, D., Parchmann, I. & Schecker, H. (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung*. (S. 257- 267). Berlin/Heidelberg: Springer-Verlag.

RIVARD, L.O.P. (1994). A review of writing to learn science: Implications for practice and research. *Journal of Research in Science Teaching*, 31(9), 969-983.

ROYSTON, P. (1995). A remark on Algorithm AS 181: The W test for normality. *Applied Statistics*, 44, 547–551.

RÜEDE, C. & WEBER, C. (2012). Schülerprotokolle aus unterschiedlichen Perspektiven Lesen – eine explorative Studie. *Journal für Mathematik-Didaktik*, 33(1), 1-28.

SANDMANN, A. (2014). Lautes Denken- die Analyse von Denk-, Lern- und Problemlöseprozessen. In Krüger, D., Parchmann, I. & Schecker, H. (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung*. (S. 179- 188). Berlin/Heidelberg: Springer-Verlag.

SANTA, C. M. & HAVENS, L. T. (1991). Learning through writing. In Santa, C. M. & Alvermann, D. E. (Hrsg.), *Science learning: Processes and applications*, 122-133.

SCHARFENBERG, F. (2005). *Experimenteller Biologieunterricht zu Aspekten der Gentechnik im Lernort Labor: empirische Untersuchung zu Akzeptanz, Wissenserwerb und Interesse*. Verfügbar unter www.pflanzenphysiologie.uni-bayreuth.de/didaktik-bio/en/pub/html/31120diss_Scharfenberg.pdf letzter Zugriff [19.7.2015].

SCHECKER, H. & PARCHMANN, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 45-66.

SCHEID, J. (2013). *Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: Ein Beitrag zur Entwicklung der Aufgabenkultur*. Berlin: Logos-Verlag.

SCHNOTZ, W. (1993). Mentale Repräsentationen beim Sprachverstehen. *Zeitschrift für Psychologie*, 201(2), 237-259.

SCHNOTZ, W. (1994). *Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten*. Weinheim: Beltz.

SCHNOTZ, W. (2005). An Integrated Model of Text and Picture Comprehension. In Mayer, R. (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 49-70). Cambridge: Cambridge University Press.

SCHNOTZ, W., BAADTE, C., MÜLLER, A. & RASCH, R. (2011). Kreatives Denken und Problemlösen mit bildlichen und beschreibenden Repräsentationen. In Sachs-Hombach, K., Totzke, R. (Hrsg.), *Bilder-Sehen-Denken* (S.204-252). Köln: Herbert von Halem Verlag.

SCHREIBER, N., THEYßEN, H., & SCHECKER, H. (2009). Experimentelle Kompetenz messen?!. *PhyDid A-Physik und Didaktik in Schule und Hochschule*, 3(8), 92-101.

SCHREIBER, N. & THEYSEN, H. (2015). Experimentelle Fähigkeiten unterstützt durch Schüler-selbstbeurteilungen diagnostizieren? In Bernholt, S. (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht, Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bremen 2014* (S. 654 - 656). Kiel: IPN.

SCHUMACHER, S. & ROTH, J. (2014). Darstellungskompetenz – Ein Schlüssel zum forschenden Lernen?! In Roth, J. & Ames, J. (Hrsg.), *Beiträge zum Mathematikunterricht 2014* (S. 1123–1126). Münster: WTM-Verlag.

STEINHOFF, T. (2003): Wie entwickelt sich die wissenschaftliche Textkompetenz?. *Der Deutschunterricht*, 55(3), 38-47.

STEPHANY, S., LINNEMANN, M., & BECKER-MROTZEK, M. (2013). Schreiben als Mittel des mathematischen Lernens. In Becker-Mrotzek, M., Schramm, K., Thürmann, E. & Vollmer, H. J. (Hrsg.), *Sprache im Fach – Sprachliches und fachliches Lernen* (S. 203-222). Münster: Waxmann.

- STORK, H. (1993). Sprache im naturwissenschaftlichen Unterricht. In Duit, R. & Gräber, W. (Hrsg.), *Kognitive Entwicklung und Lernen der Naturwissenschaften* (S.63-84), Kiel: IPN.
- SUMFLETH, E., KOBOW, I., TAUNALI, N., WALPUSKI, M. (2013). Fachkommunikation im Chemieunterricht. In Becker-Mrotzek, M., Schramm, K., Thürmann, E. & Vollmer, H. J. (Hrsg.), *Sprache im Fach – Sprachliches und fachliches Lernen* (S. 255-276). Münster: Waxmann.
- SÜSELBECK, G. (2003). Immer diese Protokolle. *Grundschule*, 6(35), 59.
- THEYßEN, H. (2014). Methodik von Vergleichsstudien zur Wirkung von Unterrichtsmedien. In Krüger, D., Parchmann, I., Schecker, H. (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S.67-80). Berlin: Springer Verlag.
- TIEMANN, R., KÖRBS, C. (2014). Die Fragenbogenmethode, ein Klassiker der empirischen didaktischen Forschung. In Krüger, D., Parchmann, I., Schecker, H. (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S.283-296). Berlin: Springer Verlag.
- WALLACE C.S., HAND B.B. & PRAIN V. (2004). *Writing and Learning in the Science Classroom*. Springer Science & Business Media.
- WALPUSKI, M. (2006). *Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback: eine empirische Studie*. Berlin: Logos-Verlag.
- WELLNITZ, N., FISCHER, H.E., KAUERTZ, A., MAYER, J., NEUMANN, I., PANT, H.A., SUMFLETH, E. & WALPUSKI, M. (2012). Evaluation der Bildungsstandards – eine fächerübergreifende Testkonzeption für den Kompetenzbereich Erkenntnisgewinnung. *Zeitschrift für Didaktik der Naturwissenschaft*, 18, 261-291.
- WALPUSKI, M. & SUMFLETH, E. (2007). Strukturierungshilfen und Feedback zur Unterstützung experimenteller Kleingruppenarbeit im Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 181-198.
- WITTECK, T. & EIKS, I. (2004). Versuchsprotokolle kooperativ erstellen. *Naturwissenschaften im Unterricht: Chemie*, 82-83. 54-56.
- WITZEL, A. (1982). *Verfahren der qualitativen Sozialforschung: Überblick und Alternativen*. Frankfurt: Campus-Verlag.
- YORE, L. D. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *International Journal of Science Education*, 25(6), 689–725.

Danksagung

Niemals hätte ich erwartet, dass diese Arbeit tatsächlich entsteht und sogar fertig wird. Da ich es ohne die Unterstützung von vielen, vielen Menschen nicht geschafft hätte, möchte ich euch von ganzem Herzen danken!

Ohne Prof. Dr. Björn Risch hätte ich die Promotion und die Arbeit als wissenschaftlicher Mitarbeiter nicht in Erwägung gezogen. Björn, danke dass du mich zur Promotion ermutigt hast! Danke für die Unterstützung, die endlosen Korrekturen und für deine vielen Uni-Weisheiten. Das familiäre Arbeitsumfeld macht es uns allen sehr einfach auch mal einen anstrengenden Tag zu überstehen.

Prof. Dr. Jürgen Roth danke ich vielmals für die zielstrebige Unterstützung während der Forschungsinitiative II, dem Schreiben der gemeinsamen Publikation und für die Übernahme des Zeitgutachtens.

Der Arbeitsgruppe Chemiedidaktik der Universität Koblenz-Landau möchte ich für die Hilfe und Mitwirkung in allen Bereichen danken. Danke Marie Schehl, dass du auch noch den kleinsten Fehler auf meinem Poster gefunden hast und dir mein Leid über die Protokolle angehört hast. Danke dem Laborteam, insbesondere Helma Ederer, dass ihr mir bei der Konzeption der Einheiten mit eurer Erfahrung immer weiterhelfen konntet. Danke den Hilfskräften Imke Freudenmacher, Dominik Föhrenbach, Kirstin Noll, Alexandra Schmelzer, Andrea Rohe und Jessica Gau, die mir bei der Konzeption, dem stundenlangen „Kistenpacken“, bei der Durchführung des Projektes und sogar beim Rating mit stets guter Laune zur Seite standen. Danke an die Abschlusskandidaten Alexander Engl, Kirstin Noll und Vanessa Schad deren Masterarbeiten mich enorm weitergebracht haben.

Ich danke der ganzen Forschergemeinschaft der Forschungsinitiative II „Bildungswissenschaften“. Gerade zu Beginn meiner Promotion hat mir die Zusammenarbeit in dieser Gruppe notwendige Ideen und Strukturen gebracht. Vielen Dank Kerstin Sitter und Stefan Schumacher für die gemeinsamen Stunden des „Protokolle-Durchschauens“ und des anschließenden „Verzweifelt-Seins“.

Danke an die Kooperationsschulen in Landau und Neustadt. Nur durch Ihre Zusammenarbeit konnten das Projekt realisiert werden. Danke für die tolle Organisation und den nicht immer so einfachen Eingriff in den Schulalltag.

Ein großer Dank geht an das Methodenzentrum in Landau vor allem an Dr. Walter Schreiber. Danke Walter, ohne deine Unterstützung hätte ich den Statistikberg niemals bewältigt! Am Ende hat es sogar tatsächlich Spaß gemacht.

Ich danke natürlich auch meinen Freunden und meiner Familie, die nicht nur einmal meine Laune abfangen mussten und mich auch selten zu Gesicht bekamen.

Und zu guter Letzt danke ich meinem Mann Alexander Engl. Danke, dass du dir bei der Arbeit ein Büro mit mir geteilt hast und dadurch viele kleine Probleme direkt aufgefangen hast. Danke, dass du dir zusätzlich noch zu Hause zu jeder Tages- und Nachtzeit meine Fragen, Probleme aber auch Freude angehört und zusätzlich noch gelöst hast.

Abbildungsverzeichnis

Abbildung 1: Vorlage für ein Versuchsprotokoll aus einem Schulbuch. Quelle: PRISMA CHEMIE 7-10, 2006.	8
Abbildung 2: Beispiel für ein Verlaufsprotokoll. Quelle: EMDEN & SUMFLETH, 2012.	10
Abbildung 3: Kontrollfragen zur Entwicklung von Leitfadenprotokollen. Quelle: KRAUS & STEHLIK, 2008.	11
Abbildung 4: Anleitung zur Gestaltung einer Foto-Story. Quelle: Prechtl, 2008.	16
Abbildung 5: Hilfeblatt zur Erstellung eines Versuchsprotokolls. Quelle: HOPPE ET AL., 2013.	17
Abbildung 6: Schülerprotokoll aus dem Schülerlabor Nawi-Werkstatt der Universität Koblenz-Landau.	19
Abbildung 7: Fragenkatalog der Fragebogenstudie zum Protokollieren in Schülerlaboren.	20
Abbildung 8: Relative Häufigkeiten der Antworten auf das Item „Wie werden die Protokollmöglichkeiten von Schülern der Klassenstufe X Ihrer Einschätzung nach genutzt?“ unterteilt nach Altersstufe.	21
Abbildung 9: Operationalisierung des Kompetenzbereichs Erkenntnisgewinnung und Einordnung des Erkenntnisprozesses.	24
Abbildung 10: Strukturierung der Teilprozesse des Erkenntnisweges nach unterschiedlichen Autoren. Quelle: modifiziert nach EMDEN (2011).	26
Abbildung 11: Angepasste Struktur der Teilschritte des Erkenntnisprozesses.	27
Abbildung 12: Zeitlicher Verlauf der empirischen Studie.	34
Abbildung 13: Übersicht der Einteilung in die Untersuchungsgruppen; Experimentalgruppe (EG), Kontrollgruppe (KG).	35
Abbildung 14: Übersicht des Vergleichsgruppendesigns, BeginnN = Beginn Nawi-Werkstatt, BeginnM = Beginn „Mathe ist mehr“.	36
Abbildung 15: Aufforderung im Arbeitsheft zum Protokollieren.	36
Abbildung 16: Anbringung und Unterteilung der Hilfsmittel in der Experimentalgruppe.	37
Abbildung 17: Leitfragen des problemzentrierten Interviews.	40
Abbildung 18: Checkliste als Hilfestellung für die Versuchsdurchführung.	43
Abbildung 19: Beispiel für eine Kontrollfrage nach dem Experiment. Bei richtiger Antwort, kommt man zur richtigen Box.	43
Abbildung 20: Gemittelte Wortanzahl der Protokolle mit Standardfehler im Vergleich der sieben Experimente. Die Stichprobengröße $n = 63$ berechnet sich aus 441 Protokollen geteilt durch 7 Experimente.	52
Abbildung 21: Wortanzahl der Protokolle aus den Untersuchungsgruppen EG und KG.	53
Abbildung 22: Wortanzahl der Protokolle in Abhängigkeit der Repräsentationsform der Protokolle.	54
Abbildung 23: Absolute Häufigkeit der gewählten Repräsentationsform in Abhängigkeit der Untersuchungsgruppe EG und KG mit dem Durchschnitt beider Untersuchungsgruppen als Vergleich.	55
Abbildung 24: Relative und absolute Häufigkeit der Verwendung des Protokollierbereichs pro Experiment.	55
Abbildung 25: Wortanzahl der Protokolle der Experimente drei und vier aus den Gruppen BeginnN und BeginnM.	57
Abbildung 26: Wortanzahl der Protokolle aus Experiment drei und vier in Bezug auf die Nawi Schulnote.	58
Abbildung 27: Häufigkeit Protokollieren und Experimentieren in Abhängigkeit der Gruppenzugehörigkeit BeginnM und BeginnN mit dem Durchschnitt beider Gruppen als Vergleich.	60
Abbildung 28: Sozialformen beim Protokollieren und Experimentieren, identische Protokollierhilfe sowie Absprache bei der Wahl der Protokollierhilfe dargestellt als Mittelwerte mit Standardfehlern.	61
Abbildung 29: Zeitlicher Verlauf der empirischen Studie.	74

Abbildung 30: Einteilung der Vergleichsgruppen.....	74
Abbildung 31: Zeitlicher Verlauf pro Experiment und Interventionsmaßnahme.	75
Abbildung 32: Gestaltung der Arbeitshefte: links die Seite aus dem Heft der Experimentalgruppe, rechts die Seite aus dem Heft der Kontrollgruppe.	76
Abbildung 33: Schülergruppe prüft die Inhaltsstoffe der Brause auf ihre Eigenschaften.....	87
Abbildung 34: Schülergruppe beim Einleiten des unbekanntes Gases in Bromthymolblau.	88
Abbildung 35: Löslichkeitsverhalten von Kohlenstoffdioxid, Ausschnitt aus der Präsentation zur Einführung. ...	89
Abbildung 36: Bedeutung der Messwerte im Messzylinder beim Löslichkeitsverhalten von Kohlenstoffdioxid.	90
Abbildung 37: Schülergruppe beim Untersuchen des unbekanntes Pulvers.	91
Abbildung 38: Schülergruppe beim Ablesen der Temperatur der Atmosphäre im Kolben.	92
Abbildung 39: Vorgehen bei der Bewertung von Protokollen. Links ein Erarbeitungsprotokoll, rechts ein Ausschnitt aus dem Kodierplan der Kategorie Produkt.	100
Abbildung 40: Lernerfolg: Auswertung von Experiment 2 über die drei Messzeitpunkte im Vergleich Kontroll- und Experimentalgruppe.	112
Abbildung 41: Lernzuwachs: Auswertung von Experiment 2 in einer Punktedifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.....	113
Abbildung 42: Lernerfolg: Auswertung von Experiment 3 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.	114
Abbildung 43: Lernzuwachs: Auswertung von Experiment 3 in einer Punktedifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.....	114
Abbildung 44: Lernerfolg: Auswertung von Experiment 5 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.	115
Abbildung 45: Lernzuwachs: Auswertung von Experiment 5 in einer Punktedifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.....	116
Abbildung 46: Lernerfolg: Auswertung von Experiment 6 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.	116
Abbildung 47: Lernzuwachs: Auswertung von Experiment 6 in einer Punktedifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.....	117
Abbildung 48: Lernerfolg: Auswertung der Experimente 2, 3, 5 und 6 über die drei Messzeitpunkte im Vergleich von Kontroll- und Experimentalgruppe.....	118
Abbildung 49: Lernzuwachs: Auswertung der Experimente 2, 3, 5 und 6 in einer Punktedifferenz von Pre- zu Post-Test und von Post- zu Follow-Up-Test im Vergleich von Kontroll- und Experimentalgruppe.	119
Abbildung 50: Protokolle zu Experiment 2 der ID-Nummern 257207 und 254109.	122
Abbildung 51: Kleingruppe beim Experiment.	127

Tabellenverzeichnis

Tabelle 1: Definition der verschiedenen Protokollformen, deren Einsatzmöglichkeiten und die verwendete Literatur	6
Tabelle 2: Vergleich der Strukturen eines Experimentes und des dazugehörigen Protokollinhaltes	7
Tabelle 3: Mögliche Unterstützungsmaßnahmen beim Protokollieren	15
Tabelle 4: Operationalisierung des Kompetenzbereich Kommunikation in Kompetenzteilbereiche und Aspekte. Quelle: KOBOW (2015).....	28
Tabelle 5: Einsatz der Erhebungsinstrumente zu den jeweiligen Untersuchungszeitpunkten	38
Tabelle 6: Übersicht der eingesetzten Experimente im Verlauf der Intervention	42
Tabelle 7: Kategorien für die Bewertung aller Forscherhefte	48
Tabelle 8: Kategorien für die Bewertung der Videoaufnahmen	50
Tabelle 9: Häufigkeitsverteilung der Nennung der Antwortmöglichkeit eines Items für Experiment 2	78
Tabelle 10: Häufigkeitsverteilung der Nennung der Antwortmöglichkeit eines Items für Experiment 6	78
Tabelle 11: Häufigkeitsverteilung der Nennung der Antwortmöglichkeit eines Items für Experiment 3	79
Tabelle 12: Einsatz der Erhebungsinstrumente zu den jeweiligen Untersuchungszeitpunkten	80
Tabelle 13: Übersicht über die eingesetzten Experimente im Verlauf der Intervention	86
Tabelle 14: Kategorien und beinhaltetete Aspekte der Auswertung	99
Tabelle 15: Reliabilitätskoeffizient für zwei Videoitems	104
Tabelle 16: Validitätskoeffizient der Bewertungen der naturwissenschaftlichen Befragung und dem Erarbeitungsprotokoll	106
Tabelle 17: Varianzanalyse einzelner Variablen in Abhängigkeit zur Zugehörigkeit zu den Interventionsgruppen.....	108
Tabelle 18: Lernerfolg: Auswertung des Fachtests zu den einzelnen Messzeitpunkten.....	109
Tabelle 19: Interaktionseffekte der Kontrollvariablen auf den Lernerfolg der Experimente	110
Tabelle 20: Interaktionseffekte der Kontrollvariablen auf den Lernzuwachs von Pre- zu Posttest der sechs Experimente.....	111
Tabelle 21: Ergebnisse der Varianzanalyse der abhängigen Variable Protokollqualität mit unabhängigen Variablen	121

Anhang

- I. Fragebogen Schülerlabore (Teilstudie 1)
- II. Kodiermanual zu Auswertung der Forscherhefte (Teilstudie 2)
- III. Kodiermanual zu Auswertung der Videoaufnahmen (Teilstudie 2)
- IV. Kodiermanual zu Auswertung der Interviews (Teilstudie 2)
- V. Testheft (Pre-Test) (Teilstudie 3)
- VI. Arbeitshefte mit sechs Experimentieranleitungen (Teilstudie 3)
- VII. Adaptiertes Kodiermanual zur Protokollauswertung von Experiment 2 „Welches Gas entsteht beim Sprudeln der Brause?“ (Teilstudie 3)

I. Fragebogen Schülerlabore (Teilstunde 1)

1. Welche Altersklassen besuchen Ihren außerschulischen Lernort?
 - Grundschüler
 - Orientierungsstufe (5. und 6. Klasse)
 - Mittelstufe (7.-10. Klasse)
 - Oberstufe
 - Studenten
 - Sonstige _____

Unterfragen je nach Altersgruppe

Wie werden die Protokollmöglichkeiten von Schülern der Grundschule Ihrer Einschätzung nach genutzt?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Sehr genutzt | Gut | angemessen | wenig |

Wie werden die Protokollmöglichkeiten von Schülern der 5. und 6. Klasse Ihrer Einschätzung nach genutzt?

Sehr genutzt Gut angemessen wenig

Wie werden die Protokollmöglichkeiten von Schülern der 7.- 10. Klasse Ihrer Einschätzung nach genutzt?

Sehr genutzt Gut angemessen wenig

Wie werden die Protokollmöglichkeiten von Schülern der Oberstufe Ihrer Einschätzung nach genutzt?

Sehr genutzt Gut angemessen wenig

Wie werden die Protokollmöglichkeiten von Studenten Ihrer Einschätzung nach genutzt?

Sehr genutzt Gut angemessen wenig

Wie werden die Protokollmöglichkeiten von Teilnehmern dieser Altersklasse Ihrer Einschätzung nach genutzt?

Sehr genutzt Gut angemessen wenig

2. Welche Materialien benutzen die Teilnehmer?

- Laborheft und Stift
- Arbeitsblätter mit Vorgaben zum Beispiel Tabellen
- Laborheft und Arbeitsblätter
- Eigene Materialien
- Keine
- Sonstiges _____

3. Fordern Sie die Teilnehmer explizit zum Protokollieren auf?
- Ja, das Protokollieren ist in den Ablauf eingebunden.
 - Ja, zu Beginn werden alle Teilnehmer zum Protokollieren aufgefordert.
 - Nein, das muss jeder Teilnehmer selbstständig entscheiden.
 - Nein, protokollieren ist uns nicht wichtig.
 - Sonstiges _____

II. Kodiermanual zu Auswertung der Forscherhefte (Teilstudie 2)

Kategorie	Code	Ausprägung
Wortanzahl <i>„Wie viele Wörter wurden geschrieben?“</i>	0-x	„ <i>Wort</i> “: Zahlen mit der Einheit oder Symbole entsprechen einem Wort. Satzzeichen werden nicht als Wörter gezählt.
Verwendung des Protokollierbereiches <i>„Wurde der vorgegebene Protokollierbereich verwendet?“</i>	0 1	„ <i>nicht verwendet</i> “: Die Schüler machen weitere Anmerkungen außerhalb des vorgegebenen Protokollierbereiches „ <i>verwendet</i> “: Die Schüler schreiben ausschließlich auf dem vorgegebenen Protokollierbereich.
Repräsentationsform <i>„Welche Repräsentationsform der Protokolle wurde verwendet?“</i>	1 2 3	„ <i>Text</i> “: Die Schüler schreiben Stichpunkte oder/und zusammenhängende Sätze <u>oder</u> wählen die Protokollierhilfe „Text“ „ <i>Tabelle</i> “: Die Schüler fertigen eine Tabellen bzw. Listen an <u>oder</u> benutzen die Protokollierhilfe „Tabelle“. „ <i>Skizze</i> “: Die Schüler zeichnen eine Grafik bzw. stellen Zusammenhänge mit Hilfe von Symbolen oder Pfeilen dar <u>oder</u> verwenden die Protokollierhilfe „Skizze“.

III. Kodiermanual zu Auswertung der Videoaufnahmen (Teilstudie 2)

Kategorie	Code	Ausprägung
Protokollieren <i>„Protokollieren die Schüler?“</i>	0	<i>„Protokollieren nicht“</i> : Die Schüler protokollieren während des Experimentes nicht in ihr Arbeitsheft.
	1	<i>„Protokollieren“</i> : Die Schüler protokollieren während des Experimentes in ihr Arbeitsheft.
Sozialform Protokollieren <i>„In welcher Sozialform protokollieren die Schüler?“</i>	0	<i>„Jeder für sich“</i> : Mindestens eine Person spricht sich nicht beim Protokollieren mit den anderen Gruppenmitgliedern ab.
	1	<i>„Zusammen“</i> : Die ganze Gruppe spricht sich beim Protokollieren ab.
Experimentieren <i>„Experimentieren die Schüler?“</i>	0	<i>„Experimentieren nicht“</i> : Alle Gruppenmitglieder sind weder aktiv noch passiv (schauen nicht zum Experiment) am Experiment beteiligt. Die Gruppe experimentiert nicht.
	1	<i>„Experimentieren“</i> : Mindestens ein Gruppenmitglied ist aktiv oder passiv (beobachten) am Experimentieren.
Sozialform Experimentieren <i>„In welcher Sozialform experimentieren die Schüler?“</i>	0	<i>„Nicht gemeinsam“</i> : Mindestens eine Person ist weder aktiv noch passiv (beobachten) am Experiment beteiligt.
	1	<i>„Zusammen“</i> : Alle Gruppenmitglieder sind aktiv oder passiv am Experiment beteiligt.
Absprache Protokollierhilfe <i>„Wurde die Protokollierhilfe in Absprache ausgesucht?“</i>	0	<i>„Keine Absprache“</i> : Es erfolgt keine Absprache zum Beispiel „Hol du mal Hilfen!“ oder „Ich hab für alle jetzt die Hilfe geholt“.
	1	<i>„Absprache“</i> : Die Schüler laufen zusammen zum Schrank und holen sich Protokollierhilfen <u>oder</u> sprechen sich vorher ab, welche Protokollierhilfe mitgebracht werden soll.

IV. Kodiermanual zu Auswertung der Interviews (Teilstudie 2)

Kategorie	Code	Ausprägung mit Ankerbeispiel
Dauer <i>„Wie lange dauert das Interview?“</i>	0-x	<i>„Zeit“</i> : Ein Schüler wird interviewt über eine bestimmte Zeit in Minuten interviewt.
Vorerfahrung Protokollieren <i>„Haben die Schüler Vorerfahrungen mit dem Protokollieren?“</i>	0 1	<i>„Keine Vorerfahrung“</i> : <u>Engl</u> : „Ok. Also ihr ähm, ihr schreibt irgendwie Sachen auf, aber nicht zu einem Experiment? Also ihr führt nicht ein Experiment durch und schreibt dann irgendwie auf, was passiert ist oder so.“ <u>J.</u> : „Nein, das machen wir nicht.“ <i>„Vorerfahrung“</i> : <u>J.</u> : „Durchführung, Beobachtung“ und alles...und da müssen wir alles schreiben... Chemikalien...Da müssen wir das schon ausführlich schreiben. Und eben auch mit Fachbegriffen.“
Einstellung zum Protokollieren <i>„Welche Einstellung haben die Schüler zum Protokollieren?“</i>	0 1 2	<i>„Unbeliebt“</i> : <u>J.</u> : „Äh, Spaß macht mir’s nicht aber...ein bisschen nervig ist es schon. *grinst* äh...weil ich andauernd Stift rausholen und...manchmal tut mir auch mein Knöchel weh.“ <i>„Gleichgültig“</i> : <u>R.</u> : „Pff, nicht schlimm. Langweilig *lacht*.“ <i>„Beliebt“</i> : <u>J.</u> : „Und mitschreiben gehört dazu und ich...es macht auch manchmal Spaß richtig. Also...da zu knobeln, was essen jetzt das genau, warum wird das so...“

Kategorie	Code	Ausprägung mit Ankerbeispiel
Bedeutung des Protokollierens <i>„Sehen die Schüler eine Bedeutung im Protokollieren?“</i>	0	„Keine Relevanz“: Der Schüler sieht keine Bedeutung bzw. keine Relevanz im Protokollieren.
	1	„Relevanz“: <u>R.</u> : „Dass man es in Erinnerung behält, dass man's sich immer wieder durchlesen kann und dass man damit auch lernen kann.“
Sozialform Protokollieren <i>„In welcher Sozialform Protokollieren die Schüler?“</i>	0	„Einzelarbeit“: <u>J.</u> : „Eigentlich hat jeder eher für sich geschrieben. Er hat glaub' ich was anderes.“
	1	„Zusammenarbeit“: <u>T.</u> : „Mit dem... hab' ich die halt auch mit ihm gemacht. Bloß der Deckel hab' ich anders gemalt wie er *kichert*.“
Probleme beim Protokollieren <i>„Welche Probleme haben die Schüler beim Protokollieren?“</i>	0	„Kein Problem“: <u>J.</u> : „Nö, is mir eigentlich egal. Die Fragen hab' ich dann halt einfach so im Kopf gehabt. Weil man hat ja auch geseh'n, was da unter'm Mikroskop lag und so.“
	1	„Problem mit dem Protokollierbereich“: <u>J.</u> : „Also des Bessere war, wo auf der gleichen Seite das auch war. Ähm,...weil da eben nicht umblättern musst und erst die Frage lesen musst.“
	2	„Problem mit der Zeit“: <u>M.</u> : „Mmhmm. Habs nicht mehr ganz geschafft, weil die Zeit nicht mehr so doll gereicht hat.“
	3	„Problem mit der Protokollierhilfe“: <u>F.</u> : „Mmhmm. Des da hab' ich nicht so ganz kapiert, mit dem, wie wir die...Werte aufschreiben sollen...Ja, ob man da die Zahlen hinschreiben soll oder die Farben.“
	4	„Sonstiges Problem“: Der Schüler äußert sich über weitere Probleme, beispielsweise das hohe Anforderungsniveau, Unterbrechungen oder Unübersichtlichkeit.

V. Testheft (Pre-Test) (Teilstudie 3)

Liebe Schülerin, lieber Schüler,

wir haben dir heute ein Heft mit verschiedenen Aufgaben und Fragen mitgebracht.

Bevor du anfängst, es auszufüllen, lies dir bitte alle Anleitungen und Fragen genau durch.

Bitte schreibe nicht von deinem Nachbarn ab! Es ist wichtig, dass du die Fragen so beantwortest, wie du es für richtig hältst.

Wenn du dir bei einer Antwort nicht sicher bist, dann antworte so, wie es dir am ehesten richtig erscheint.

Vielen Dank für deine Mitarbeit!

Jetzt kann es losgehen!

1 Zunächst haben wir ein paar allgemeine Fragen über dich.

Deine Schüler ID: _____

Tag deines Geburtstags (in Zahlen):_____

(Beispiel: Wenn du am 05. März Geburtstag hast, schreibst du „05“.)

Wie alt bist du?

_____ Jahre

Du bist...

ein Junge

ein Mädchen

Zuhause sprichst du...

immer Deutsch

manchmal eine andere Sprache

immer eine andere Sprache

Wo sind deine Eltern geboren?

Mutter Deutschland

anderes Land

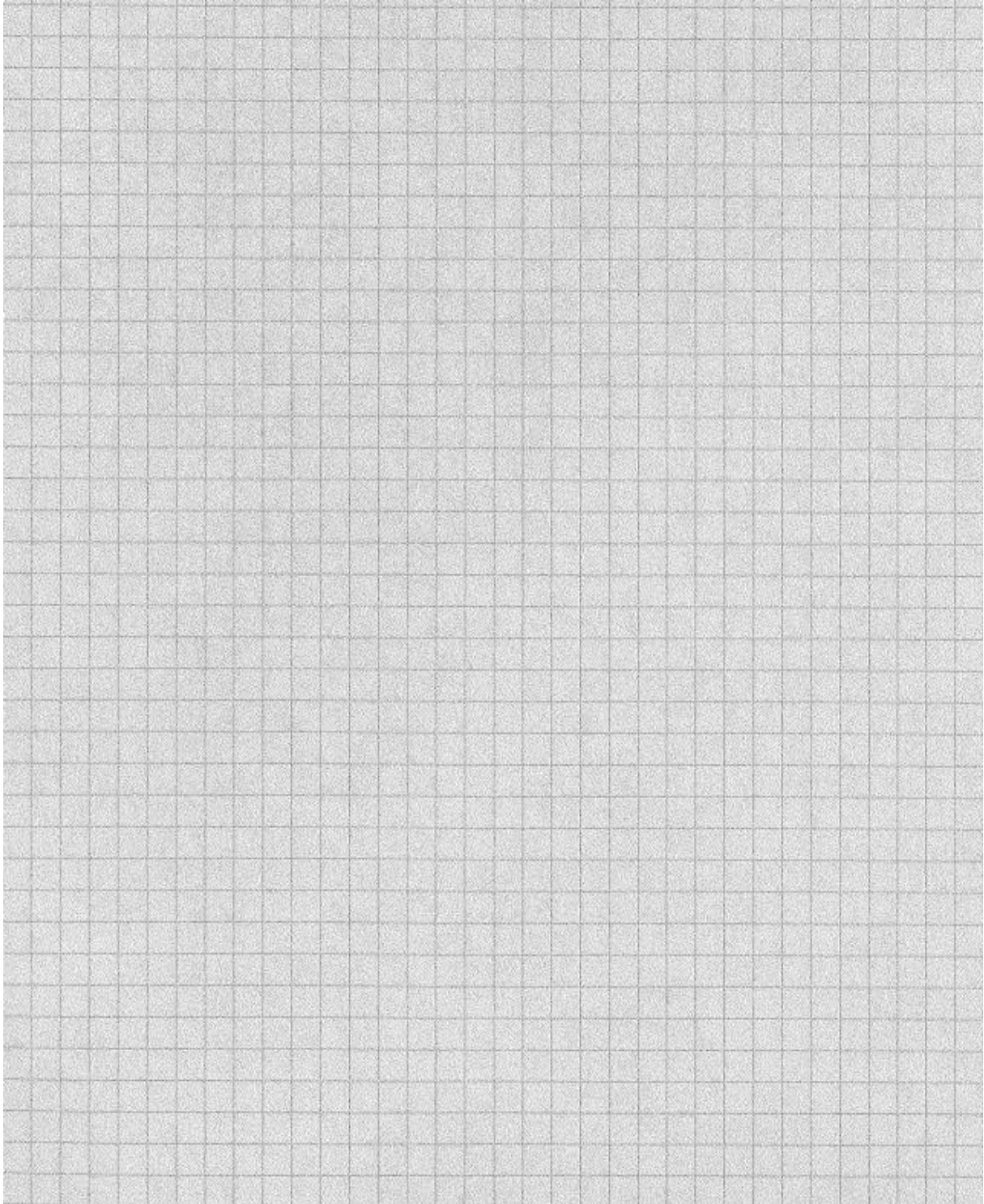
Vater Deutschland

anderes Land

2 Videoaufgabe

Für die nachfolgende Aufgabe hast du insgesamt 7 Minuten Bearbeitungszeit.

Schau genau hin, was du siehst. Du siehst gleich einen Film über ein Experiment (Länge: 2,50 Minuten). Schreibe und zeichne alles auf, was du beobachtet hast. Notiere es so, dass du anderen Kindern das Experiment erklären könntest.



3 Eine kleine Befragung

Nun wollen wir wissen, wie deine Meinung **zu dem Fach Naturwissenschaften (Nawi)** ist. Bitte lies die folgenden Sätze genau durch, und kreuze die Antwort an, die im Allgemeinen am besten auf dich passt.

Zum Beispiel: Wenn die folgende Aussage für dich „**eher nicht stimmt**“, kreuzt du so an:

	stimmt voll und ganz	stimmt fast ganz	stimmt eher	stimmt eher nicht	stimmt fast gar nicht	stimmt gar nicht
1. Ich mag Naturwissenschaften.				✗		

Wichtig: Triff ehrliche Entscheidungen! Dieser Fragebogen hat keinen Einfluss auf deine Note. Dein Lehrer erfährt deine Antworten nicht. Überlege bei den Aussagen nicht lange. Es gibt keine richtigen oder falschen Lösungen. Die Antwort ist "richtig", die am besten deiner Meinung entspricht. Kreuze für jede Aussage immer nur eine Möglichkeit an, und lass bitte keine Aussage unbeantwortet. Frage bitte nach, wenn dir etwas unklar ist!

	stimmt voll und ganz	stimmt fast ganz	stimmt eher	stimmt eher nicht	stimmt fast gar nicht	stimmt gar nicht
1. Ich finde Naturwissenschaften sehr interessant.						
2. Ich bin immer sehr gespannt darauf, wie gut ich in Naturwissenschaften abschneide.						
3. Mir fällt Naturwissenschaften leicht.						
4. Ich fühle mich durch Aufgaben in Naturwissenschaften herausgefordert.						
5. Naturwissenschaften macht mir viel Spaß.						
6. Ich bin immer fest entschlossen, mich in Naturwissenschaften voll anzustrengen.						
7. Ich mag Naturwissenschaften.						
8. Ich bin für Naturwissenschaften begabt.						

4 Fragen zur Nawi-Werkstatt

In der Nawi-Werkstatt werdet ihr verschiedene Experimente bearbeiten. Wir würden gerne wissen, ob ihr einige Experimente vorher schon kennt. Einige dieser Themen habt ihr in der Schule noch nicht behandelt, deshalb ist es auch nicht schlimm, wenn ihr die Antwort nicht wisst.

Kreuze die richtige Antwort an. Es gibt immer nur **eine** richtige Antwort! Bei freien Feldern musst du eine Antwort hin schreiben.

1. In einem Experiment wird untersucht, welche Kombination der vier Stoffe Zucker, Weinsäure, Natriumhydrogencarbonat und Aroma bei Wasserzugabe sprudelt. Nenne einen Inhaltsstoff bei dem es auch zusammen mit Weinsäure **nicht sprudelt**, wenn man Wasser dazu gibt?

2. Zucker und Aroma sind in der Brause....

- nicht für das Sprudeln verantwortlich.
- nicht für den Geschmack verantwortlich.
- für das Sprudeln verantwortlich.
- für nichts verantwortlich.

3. Welche zwei Voraussetzungen sind wichtig, um die Brause zum Sprudeln zu bekommen?

- Es müssen die Stoffe Natriumhydrogencarbonat und Aroma vorhanden sein. Anschließend wird Wasser hinzu gegeben.
- Es müssen die Stoffe Natriumhydrogencarbonat und Weinsäure vorhanden sein. Wasser wird nicht benötigt.
- Es müssen die Stoffe Zucker und Aroma vorhanden sein. Anschließend wird Wasser hinzu gegeben.
- Es müssen die Stoffe Natriumhydrogencarbonat und Weinsäure vorhanden sein. Anschließend wird Wasser hinzu gegeben.

4. Löst man Luft in Wasser und gibt Bromthymolblau hinzu, hat die Lösung die Farbe ...
- farblos
 - gelb
 - grün
 - blau
5. In einem Experiment werden Weinsäure und Natriumhydrogencarbonat miteinander gemischt und Wasser hinzu gegeben. Das entstehende Gas wird in eine Bromthymolblau-Lösung eingeleitet. Wie verändert sich die Farbe?
- Farblos zu gelb.
 - Gelb zu blau.
 - Grün zu blau.
 - Blau zu gelb.
6. In einem Experiment werden die Gase Luft, Kohlenstoffdioxid und Distickstoffoxid miteinander verglichen. Welche Eigenschaften treffen auf Kohlenstoffdioxid zu?
- Löscht die Kerzenflamme und die Bromthymolblau-Lösung wird gelb
 - Löscht die Kerzenflamme nicht und die Bromthymolblau-Lösung wird gelb
 - Löscht die Kerzenflamme nicht und die Bromthymolblau-Lösung bleibt blau
 - Löscht die Kerzenflamme und die Bromthymolblau-Lösung bleibt blau
7. Wie viele Inhaltsstoffe sind für das Sprudeln von Brause notwendig?
- 0
 - 3
 - 1
 - 2

8. Wie wird das entstandene Gasvolumen von Brausetabletten in Wasser gemessen?

- Brausetablette in einen Messzylinder legen und mit Wasser befeuchten.
- Umgedrehten Messzylinder in einer Plastikwanne mit Luft befüllen und Brausetablette darunter legen.
- Umgedrehten Messzylinder in einer Plastikwanne mit Wasser befüllen und Brausetablette darunter legen.
- Brausetablette unter einen Luftballon legen und mit Wasser befeuchten.

9. In einem Experiment wird eine unbekannte Brausetablette bestimmt indem das entstehende Gasvolumen von drei bekannten Brausetabletten mit dem Gasvolumen der unbekanntes Brausetablette verglichen wird. Wie könnte beispielsweise dein Ergebnis aussehen, wenn Brausetablette B die unbekanntes Brausetablette wäre?

Brausetablette A: _____ mL Gas
Brausetablette B: _____ mL Gas
Brausetablette C: _____ mL Gas
Unbekanntes Brausetablette: _____ mL Gas

10. In einem Experiment werden Traubenzucker, Salz und Zitronensäure in einem Löffel über einer Kerze erhitzt. Welches Pulver schmilzt **nicht**, wenn man es über einer Kerze erhitzt?

- Traubenzucker
- Salz
- Zitronensäure
- Alle Pulver schmelzen

11. In einem Experiment werden Traubenzucker, Salz und Zitronensäure in eine Bromthymolblau-Lösung gegeben. Bei welchem Pulver ändert Bromthymolblau seine Farbe?

- Traubenzucker
- Salz
- Zitronensäure
- Bei keinem Pulver verändert sich die Farbe

12. In einem Experiment wird eine Brausetablette im normalen (etwa 18°C), sauren (etwa 18°C) und warmen (etwa 60°C) Wasser gelöst. Bei welcher Bedingung wird **im Messzylinder** das wenigste Kohlenstoffdioxid gemessen?

13. In einem Experiment wird eine Brausetablette im normalen (etwa 18°C), sauren und warmen (etwa 60°C) Wasser gelöst. Welche Aussage ist richtig?

Wenn man die Bedingungen des Wassers ändert...

- löst sich immer mehr Kohlenstoffdioxid im Wasser.
- löst sich immer weniger Kohlenstoffdioxid im Wasser.
- löst sich immer gleich viel Kohlenstoffdioxid im Wasser.
- löst sich je nach Bedingung unterschiedlich viel Kohlenstoffdioxid im Wasser.

14. In einem Experiment werden die Gase Luft und Kohlenstoffdioxid durch Lichteinstrahlung erwärmt. Bei welchem Gas steigt die Temperatur mehr an?

- Kohlenstoffdioxid
- Bei beiden Gasen steigt die Temperatur gleich.
- Luft
- Die Temperatur steigt nicht an.

15. Welches Gas entsteht, wenn Weinsäure und Natriumhydrogencarbonat miteinander gemischt und Wasser hinzu gegeben wird?

16. Welche Schlussfolgerung ist richtig?

Ein unbekanntes Gas kann bestimmt werden, indem...

- Eigenschaften des Gases durch Experimente herausgefunden werden und diese mit den Steckbriefen zu den Gasen verglichen werden.
- im Internet recherchiert wird.
- die Steckbriefe der Gase miteinander verglichen und abgeschrieben werden.
- Keine der oben genannten Schlussfolgerungen sind richtig.

17. Ein unbekanntes Gas hat die Eigenschaft Feuer zu löschen. Es könnte Luft, Distickstoffoxid oder Kohlenstoffdioxid sein. Um welches Gas handelt es sich?

- Keines der drei
- Kohlenstoffdioxid
- Luft
- Distickstoffoxid

18. In einem Experiment wird das entstehende Gasvolumen von unterschiedlichen Brausetabletten mehrerer Marken bestimmt. Jede Brausetablette entwickelt ein bestimmtes Gasvolumen. Wovon hängt das entstehende Gasvolumen einer Brausetablette ab?

- Es hängt von dem Experiment ab.
- Es hängt von der Masse an Weinsäure und Natriumhydrogencarbonat ab.
- Es hängt von der Tagesform des Experimentators ab.
- Es hängt von der Lufttemperatur ab.

19. In einem Experiment wird eine unbekannte Brausetablette bestimmt indem das entstehende Gasvolumen von den bekannten Brausetabletten mit dem Gasvolumen der unbekannt Brausetablette verglichen wird.

Brausetablette A: 56mL

Brausetablette B: 12mL

Brausetablette C: 120 mL

Unbekannte Brausetablette: 60mL

Welche Schlussfolgerung kannst du aus dem Ergebnis ziehen?

20. In einem Experiment werden Traubenzucker, Salz und Zitronensäure jeweils in ein Becherglas mit Bromthymolblau-Lösung gegeben und in einem Löffel über der Kerze erhitzt. Welche Pulver haben die folgenden Eigenschaften: 1) Schmelzen beim Erhitzen und verändern nicht die Farbe von Bromthymolblau?

- Traubenzucker und Salz
- Salz und Mehl
- Mehl und Traubenzucker
- Zitronensäure und Traubenzucker

21. In einem Experiment soll ein unbekanntes Pulver identifiziert werden. Es ist entweder Traubenzucker, Salz oder Zitronensäure. Welche Schlussfolgerung ist richtig?

Ein unbekanntes Pulver kann bestimmen werden, indem...

- das Gasvolumen des unbekannt Pulvers bestimmt wird.
- die Eigenschaften der bekannten Pulver mit denen des unbekannt Pulvers verglichen werden.
- die Eigenschaften des unbekannt Pulvers bestimmt werden.
- das Gasvolumen aller Pulver miteinander verglichen werden.

22. In einem Experiment wird eine Brausetablette im normalen (etwa 18°C), sauren (etwa 18°C) und warmen (etwa 60°C) Wasser gelöst. Bei welcher Bedingung wird **im Messzylinder** das **meiste** Kohlenstoffdioxid gemessen?

23. In einem Experiment wird eine Brausetablette im normalen (etwa 18°C), sauren (etwa 18°C) und warmen (etwa 60°C) Wasser gelöst. Bei welcher Bedingung löst sich das **meiste** Kohlenstoffdioxid **im Wasser**?

24. In einem Experiment werden die Gase Luft und Kohlenstoffdioxid durch Lichteinstrahlung erwärmt. Bei welchem Gas steigt die Temperatur **weniger** an?

- Kohlenstoffdioxid
- Bei beiden Gasen steigt die Temperatur gleich.
- Luft
- Die Temperatur steigt nicht an.

25. In einem Experiment werden die Gase Luft und Kohlenstoffdioxid durch Lichteinstrahlung erwärmt. Wie verändert sich die Temperatur?

- Bei Kohlenstoffdioxid steigt die Temperatur, bei Luft nicht.
- Bei Luft steigt die Temperatur, bei Kohlenstoffdioxid nicht.
- Bei beiden Gasen fällt die Temperatur.
- Bei beiden Gasen steigt die Temperatur.

26. Durch welches Vorgehen kann bestimmt werden, welches Gas sich durch Lichteinstrahlung am meisten erwärmt?

- Temperaturvergleich über eine unterschiedliche Zeit
- Zeitvergleich bei der gleichen Temperatur
- Temperaturvergleich über die gleiche Zeit
- Lichteinstrahlung verändern und Temperatur messen

27. Welche Schlussfolgerung ist richtig?

Eine unbekannte Brausetablette kann aus drei bekannten Brausetabletten bestimmt werden, indem...

- die Steckbriefe der Brausetabletten verglichen werden.
- die Messwerte des entstehenden Gasvolumina miteinander verglichen werden.
- die vier Brausetabletten gewogen werden.
- die Reaktion der vier Brausetabletten im Wasser beobachtet wird.

28. In einem Experiment wird untersucht, welcher der vier Stoffe Zucker, Weinsäure, Natriumhydrogencarbonat und Aroma bei Wasserzugabe sprudelt. Bei welcher **Kombination von zwei Stoffen** sprudelt es, wenn man Wasser dazu gibt?

Vielen Dank für deine Mitarbeit

VI. Arbeitshefte mit sechs Experimentieranleitungen (Teilstudie 3)

Was bringt die Brause zum Sprudeln?



Brause besteht aus mehreren Inhaltsstoffen.

Zutaten: Zucker, Säuerungsmittel: L(+)-Weinsäure, Natriumhydrogencarbonat, Süßstoffe: Natriumcyclamat, Acesulfam-K und Saccharin-Natrium, Aroma, Farbstoffe: E160a, E160e, E163, E141.

Den Inhalt des Portionsbeutels in ein Glas mit 0,2 Liter frischem Wasser (Mineralwasser) schütten. Sofort erhält man ein herrlich prickel-frisches Getränk.

Aber welche Stoffe sind jetzt für das Sprudeln der Brause in Wasser verantwortlich und welche nicht?



1. Betrachte eine kleine Portion Brausepulver in der Petrischale unter dem Binokular.

Falls du Unterstützung bei der Bedienung des Binokulars benötigst, findest du in deiner Kiste eine Hilfefkarte.

2. Die Stoffe in den Rollrandgläsern

- Natriumhydrogencarbonat
- Zucker
- Aroma
- Weinsäure

entnimmst du mit einem Spatel aus dem Rollrandglas und gibst sie in eine Petrischale. Achte darauf, dass pro Stoff nur ein Spatel verwendet wird.

3. Zu dem Stoff gibst du etwas Wasser aus der Spritzflasche.

4. Die Stoffe werden erst alle einzeln getestet und dann miteinander kombiniert getestet.

Welches Gas entsteht beim Sprudeln der Brause?



Bei der richtigen Kombination von Zutaten sprudelt die Brause. Die Blasen im Wasser zeigen, dass ein Gas beim Sprudeln entsteht. Es gibt aber sehr viele Gase. **Welches Gas entsteht wenn man Weinsäure und Natriumhydrogencarbonat mischt und welche Eigenschaften hat es?**



Eigenschaft Brennbarkeit

1. Stelle die Teelichter auf die Tischkante und zünde sie an.
2. Gebe drei Löffel Weinsäure und drei Löffel Natriumhydrogencarbonat in die leere Flasche und anschließend etwas Wasser dazu.
3. Montiere den Deckel zügig so, dass die untere Öffnung nicht mit der Flüssigkeit in Kontakt kommt.
4. Versuche die Teelichter ausschließlich mit dem entstehenden Gas zu löschen.



Reaktion mit Bromthymolblau

1. Fülle drei Bechergläser halb mit Wasser und gebe jeweils 10 Tropfen Bromthymolblau dazu. Das Wasser wird blau.
2. Gebe drei Löffel Weinsäure und drei Löffel Natriumhydrogencarbonat in die leere Flasche und anschließend etwas Wasser dazu. Montiere den Deckel zügig so, dass die untere Öffnung nicht in die Flüssigkeit reicht.
3. Leite das entstehende Gas in das Becherglas mit Bromthymolblau ein.
Leite nun Luft (aus der Luftpumpe) und Distickstoffoxid (aus Sahnekapselfeln) mit Hilfe eines Luftballons und einem Schlauch in je ein weiteres Becherglas mit Bromthymolblau ein.



Vergleiche dein Ergebnis mit den Steckbriefen!

Die unbekannte Brausetablette

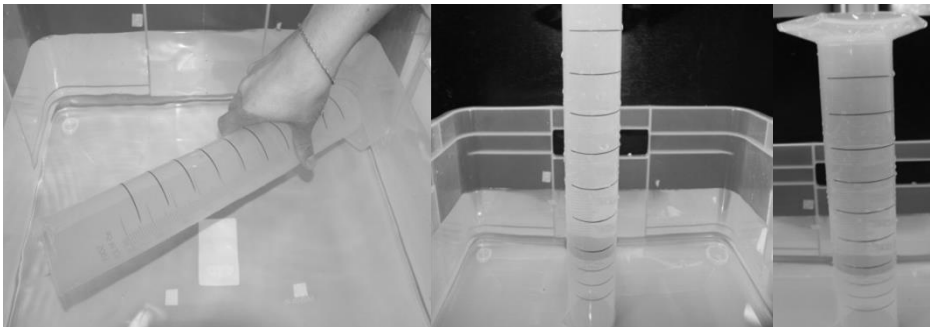


Eure Aufgabe ist es, eine unbekannte Brausetablette zu bestimmen. Es gibt aber viele unterschiedliche Tabletten von mehreren Herstellern. **Wie kann man unterschiedliche Brausetabletten identifizieren und welche ist die unbekannte Brausetablette?**



Um die unbekannte Brausetablette zu identifizieren, musst du wieder die Versuchsergebnisse der bekannten Brausetabletten mit dem Ergebnis der unbekannteren vergleichen.

1. Fülle eine Plastikwanne und fülle einen Messzylinder in der Plastikwanne mit Wasser und stelle ihn verkehrt herum in die Wanne, ohne dass der Messzylinder luftblasenfrei gefüllt ist.



2. Lege anschließend die erste Brausetablette unter die Öffnung und warte ab, wie viel Kohlenstoffdioxid entsteht. Halte dabei den Messzylinder fest, damit er nicht umfällt.
3. Bevor du die nächste Brausetablette benutzt, wechsele das Wasser in der Plastikwanne.
4. Mit den bekannten Brausetabletten verfährt du genauso, wie in Schritt 1 bis 3 beschrieben.
5. Führe nun mit der unbekannteren ebenso die Schritte 1 bis 3 durch.

Löslichkeit von Kohlenstoffdioxid



Unter welchen Bedingungen kann mehr oder weniger Kohlenstoffdioxid gelöst werden?

Löslichkeit von Kohlenstoffdioxid in normalem Wasser



1. Befülle die Plastikwanne bis zur 5 Liter Markierung mit kalten Leitungswasser.
2. Fülle den 250 mL Messzylinder in der Plastikwanne mit Wasser und stelle ihn verkehrt herum in die Plastikwanne, ohne dass Luftblasen im Messzylinder sind.
3. Lege anschließend eine Brausetablette unter die Öffnung des Messzylinders.

Löslichkeit von Kohlenstoffdioxid in saurem Wasser

1. Befülle die Plastikwanne bis zur 5 Liter Markierung mit kaltem Leitungswasser.
2. Gib 100 ml Essig in die Plastikwanne hinzu.

Führe nun wieder Schritt 2. - 3. aus dem ersten Teil durch.

Löslichkeit von Kohlenstoffdioxid in warmem Wasser

1. Befülle die Plastikwanne bis zur 4 Liter Markierung mit kaltem Leitungswasser.
2. Fülle den Wasserkocher bis zur 1 Liter Markierung und lass das Wasser aufkochen.
3. Gebe nun das kochende Wasser in die Plastikwanne hinzu.

Führe nun wieder Schritt 2. - 3. aus dem ersten Teil durch.

Unbekanntes Pulver



Eure Aufgabe ist es, ein unbekanntes Pulver zu bestimmen. Es gibt aber viele unterschiedliche weiße Pulver die alle gleich aussehen. Wie kann man unterschiedliche Pulver identifizieren und welche ist das unbekannte Pulver?



Zuerst werden die drei bekannten Pulver Salz, Traubenzucker und Zitronensäure mit zwei Experimenten untersucht. Achte darauf, dass du für jedes Pulver einen anderen Spatellöffel verwendest.

Reaktion mit Bromthymolblau

1. Drei Bechergläser werden jeweils halbvoll mit Leitungswasser und je 10 Tropfen Bromthymolblau gegeben. Das Wasser wird blau.
2. Anschließend werden in je ein Becherglas zwei Spatellöffel des zu untersuchenden Pulvers gegeben.

Erhitzen der Pulver

1. Es wird je ein halber Spatellöffel mit dem zu untersuchenden Pulver aus dem Rollrandglas entnommen und zwei Minuten (Stoppuhr) über eine brennende Kerze gehalten.
2. **Den Löffel sofort nach dem Versuch spülen!**

Unbekanntes Pulver

Führe nun die beiden Experimente mit dem unbekanntem Pulver durch.

Lichteinstrahlung bei verschiedenen Gasen



In unserer Atmosphäre befindet sich außer dem Gasgemisch Luft unter anderem auch das Gas Kohlenstoffdioxid. Wie wirken sich die verschiedenen Gase auf die Temperatur aus?

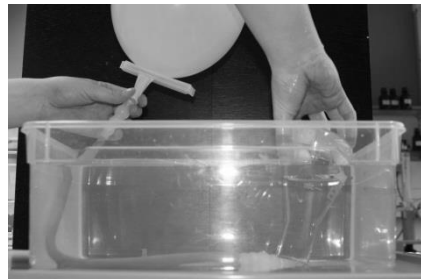
Erlenmeyerkolben nur oben anfassen!

Vorsicht! Lampe wird sehr heiß!



Luft-Atmosphäre

1. Stelle die Lampe mit ihrem Gestell auf die Markierung, aber schalt sie noch nicht ein.
2. Fülle in der Plastikwanne den Erlenmeyerkolben mit Wasser und stelle ihn verkehrt herum in die Wanne, so dass keine Luftblasen im Kolben sind.
3. Pumpe einen Luftballon mit Luft (Luftpumpe) auf und leite das Gas über den Schlauch in den Erlenmeyerkolben, so dass er vollständig gefüllt ist.
4. Der Kolben wird mit einer Petrischale verschlossen und herausgenommen. Sollte sich noch etwas Wasser im Kolben befinden, wird dieses **sehr schnell** abgelassen, so dass möglichst wenig Gas entweicht. Danach wird der Kolben außen abgetrocknet und auf die Markierung gestellt.
5. Die Petrischale wird **sehr schnell** durch einen durchbohrten Gummistopfen mit Thermometer ersetzt.
6. Drehe den Kolben so, dass die Abdeckung am Thermometer zur Lampe zeigt.



7. **WICHTIG!** Warte 2 Minuten bis die Anfangstemperatur abgelesen wird.
8. Schalte die Lampe ein und lasse von nun an die Stoppuhr laufen. Lies jede Minute die Temperatur ab.
9. Nach sechs Minuten wird das Experiment beendet und die Lampe ausgeschaltet.

Kohlenstoffdioxid-Atmosphäre

Fülle in einer Plastikwanne den Erlenmeyerkolben mit Wasser und stelle ihn auf dem Kopf.

Lege zwei Brausetabletten unter die Öffnung des Kolbens und fange das Kohlenstoffdioxid auf.



Jetzt wird das gleiche Experiment durchgeführt, wie in den Schritten 4 bis 9 beschrieben.

Die Anfangstemperatur sollte ungefähr gleich sein.

VII. Adaptiertes Kodiermanual zur Protokollauswertung von Experiment 2 „Welches Gas entsteht beim Sprudeln der Brause?“ (Teilstudie 3)

Notizen und Rechtschreibfehler sind nicht mit zu bewerten!!!

ID=6 stellig zum Beispiel 261721 (ID + Geburtstag)

VARIABLE		AUSPRÄGUNG		
Repräsentationsform Welche Repräsentationsformen sind vorhanden und stehen sie in einer Beziehung zueinander?				
R_Tabelle	Ist eine Tabelle/sind Tabellen/tabellenartiger Aufbau im Protokoll vorhanden?	0 1	nicht vorhanden vorhanden	
R_Skizze	Ist eine Skizze/sind Skizzen im Protokoll vorhanden?	0 1	nicht vorhanden vorhanden	
R_Text	Sind Stichworte/Sätze/Texte im Protokoll vorhanden?	0 1	nicht vorhanden vorhanden	
R_Fachsymbole	Sind Fachsymbole im Protokoll vorhanden?	0 1	nicht vorhanden vorhanden	
Gliederung Wert Ist das Protokoll äußerlich gegliedert? Durch welche Elemente wird es gegliedert?				
G_Vorstufe	Vorstufe: Gibt es eine ausreichende Anzahl an Elementen im Protokoll, die eine Gliederung zulassen? Skizze: mind. 2 einzelne Skizzen Text: mind. 2 Sätze bzw. Stichworte/stichwortartige Notizen	0 1	nein: nicht vorhanden ja: vorhanden	Wenn 0, dann Gliederung = 99
G_Absatz	Absatz: Gibt es Absätze/Abschnitte im Protokoll? * Wird ein fortlaufend geschriebener Text durch einen einfachen Absatz (Zeilenumbruch) unterbrochen? <u>oder</u> * Wird ein fortlaufend geschriebener Text durch einen erweiterten Absatz (Leerzeile) unterbrochen? <u>oder</u> * Werden stichwortartige Notizen durch einen Absatz unterbrochen? <u>oder</u> * Werden Einzüge (z.B. zu Beginn der ersten Zeile eines Absatzes) verwendet? <u>oder</u> * Sind Absätze zwischen den einzelnen Skizzen vorhanden? <u>oder</u> * Werden zwischen Text und Skizze Absätze gemacht? (→ rein deskriptiv, die Genauigkeit spielt keine Rolle)	0 1	nein: nicht vorhanden ja: vorhanden	0,33

Beziehung				
Gibt es eine Beziehung zwischen den Inhalten?				Wert
B_Unbek_Loescht	Das unbekannte Gas bzw. das Brausegas stehen mit dem Löschen der Kerze in Beziehung.	0 1	nicht vorhanden vorhanden	0,1
B_Unbek_Gelb	Das unbekannte Gas bzw. das Brausegas stehen mit der Färbung gelb oder „einer Veränderung“ in Beziehung.	0 1	nicht vorhanden vorhanden	0,1
B_Luft_Blau	Luft steht mit der Färbung blau oder mit „keiner Veränderung“ in Beziehung.	0 1	nicht vorhanden vorhanden	0,1
B_N2O_Blau	Distickstoffoxid steht mit der Färbung blau „keiner Veränderung“ in Beziehung.	0 1	nicht vorhanden vorhanden	0,1
B_Kohlst_Loescht	Kohlenstoffdioxid steht mit dem Löschen der Kerze in Beziehung.	0 1	nicht vorhanden vorhanden	0,2
B_Kohlst_Gelb	Kohlenstoffdioxid steht mit der Färbung gelb in Beziehung.	0 1	nicht vorhanden vorhanden	0,2
B_Unbek_Kohlst	Das unbekannte Gas bzw. das Brausegas steht mit Kohlenstoffdioxid in Beziehung.	0 1	nicht vorhanden vorhanden	0,2

Publikationen und Tagungsbeiträge

Publikationen

ENGL, L., RISCH, B. (2017). Protokollieren Trainieren – Videovignetten zur Förderung und Diagnose der Protokollierfähigkeit. *Naturwissenschaften im Unterricht Chemie*, 158, 30-33.

ENGL, L., RISCH, B. (2015). Auswirkung der Protokollierfähigkeit auf die Lernwirksamkeit. In: S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bremen 2014* (S. 265-267). Kiel: IPN.

PRETSCH, J., EHRHARDT, N., ENGL, L., RISCH, B., ROTH, J., SCHUMACHER, S., & SCHMITT, M. (2016). Injustice in school and students' emotions, well-being, and behavior: A longitudinal study. *Social Justice Research*, 29(1), 119-138.

RISCH, B., & ENGL, L. (2015). Landauer Experimentier (s) pass—Ein Schülerlabor öffnet seine Türen. *Schulpädagogik heute*, 6(11).

ENGL, L., SCHUMACHER, S., SITTER, K., GRÖBLER, M., NIEHAUS, E., RASCH, R., ROTH, J. & RISCH, B. (2015). Entwicklung eines Messinstrumentes zur Erfassung der Protokollierfähigkeit—initiiert durch Video-Items. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 223-229.

Tagungsbeiträge

ENGL, L., RISCH, B. (Vortrag GDCh 2014). Prolab - Protokollieren im Labor.

ENGL, L., RISCH, B. (Vortrag GDCh 2014). Auswirkung der Protokollierfähigkeit auf den Lernerfolg.

ENGL, L., RISCH, B. (Vortrag GDCh 2013). Bedeutung des Protokollierens beim Erfassen und Identifizieren von naturwissenschaftlichen Phänomenen am außerschulischen Lernort.

ENGL, A., SAUER, L. & RISCH, B. (Poster GDCh 2013). Einflüsse auf das Protokollierverhalten von Schülerinnen und Schülern der Orientierungsstufe am außerschulischen Lernort.

SAUER, L., RISCH, B. (Poster LELA-Jahrestagung 2013). Bedeutung des Protokollierens beim Erfassen und Identifizieren von naturwissenschaftlichen Phänomenen am außerschulischen Lernort.

SAUER, L. & RISCH, B. (Poster GDCh 2012). Bedeutung des Protokollierens am außerschulischen Lernort beim Erfassen und Identifizieren von naturwissenschaftlichen Phänomenen.