

Der Einfluss der MINIPHÄNOMENTA auf die methodisch-formalen naturwissenschaftlichen Fähigkeiten von Schülerinnen und Schülern der Primarstufe

Skizze einer Interventionsstudie im Kontext eines naturwissenschaftlichen Bildungsprojekts

Einleitung

In einem ersten Artikel zur MINIPHÄNOMENTA (vgl. widerstreit-sachunterricht, Ausgabe 11, Oktober 2008), einer Initiative zur Stärkung der naturwissenschaftlichen Grundbildung in der Primarstufe, wurden die wesentlichen konzeptionellen, lern- und bildungstheoretischen Grundlagen des Projekts herausgearbeitet. In diesem Zusammenhang konnte gezeigt werden, dass es sich bei der MINIPHÄNOMENTA nicht, wie oft angenommen, um ein Unterrichtsmedium der naturwissenschaftlichen Grundbildung in der Primarstufe handelt. Vielmehr ist die MINIPHÄNOMENTA ein Angebot an Schülerinnen und Schüler sich im Rahmen ihrer Pausenbeschäftigung mit naturwissenschaftlichen Fragestellungen auseinander zu setzen. Dies jedoch schließt eine unterrichtliche Behandlung der Versuchstationen keineswegs aus. Im Gegenteil scheint es sinnvoll zu sein die Erfahrungen der Schülerinnen und Schüler mit den Stationen der MINIPHÄNOMENTA aufzugreifen und weiter zu entwickeln.

Die MINIPHÄNOMENTA besteht aus 53 Versuchstationen, die auf den Gängen und in der Pausenhalle von Schulen der Primarstufe aufgestellt werden. Jede der Stationen stellt ein Phänomen in Form eines Versuchsaufbaus dar. Die Versuchstationen laden zum naturwissenschaftlich-technischen, vorrangig physikalischen und chemischen Experimentieren und Denken ein. In diesem Zusammenhang soll betont werden, dass die Auseinandersetzung mit den Stationen der MINIPHÄNOMENTA selbstgesteuert erfolgt. Es befinden sich folglich keinerlei Anleitungen oder Erklärungen an den Versuchstationen. Die Schülerinnen und Schüler bestimmen ihre Auseinandersetzung mit dem Phänomen frei in der Art und Weise, dass sie entscheiden ob, in welchem Umfang, wie und mit welchem Ziel sie an der Station arbeiten.

Die Versuchstationen können im Anschluss an die Teilnahme an einer Fortbildung durch die jeweilige Schule zunächst ausgeliehen werden. Die MINIPHÄNOMENTA ist aber keine Wanderausstellung. Vielmehr ist es das Ziel nach der Ausleihe Eltern für einen teilweisen Nachbau der Stationen zu gewinnen, die dann der Schule zur Verfügung gestellt werden, um so eine dauerhafte Stärkung der naturwissenschaftlichen Grundbildung in der betreffenden Schule zu erreichen. Über diese konzeptionellen Anmerkungen hinaus wurde im Rahmen des ersten Artikels der theoretische Begründungszusammenhang der MINIPHÄNOMENTA skizziert. Auf einer didaktisch-lerntheoretischen Ebene konnte zum einen ein inhaltlicher Zusammenhang mit den Annahmen Wagenscheins (vgl. hier insbesondere 1965; 1989; 1992) hergestellt werden, zum anderen wurde das Lernen der Schülerinnen und Schüler an den Stationen der MINIPHÄNOMENTA in das Rahmenmodell des selbstgesteuerten Lernens nach Schiefele und Pekrun (1996) eingeordnet. Im Rahmen einer bildungstheoretisch orientierten Diskussion konnte darüber hinaus gezeigt werden, dass das Konzept der naturwissenschaftlichen Grundbildung, verstanden als scientific literacy (vgl. Rost, Prenzel, Carstensen, Senkbeil 2004), zur bildungstheoretischen Fundierung der MINIPHÄNOMENTA nicht ausreichend ist. Ein wesentlicher Aspekt der Argumentation war dabei, dass das Lernen an den Versuchstationen der MINIPHÄNOMENTA eine deutlich erweiterte Zieldimension im Vergleich zum Konzept der scientific literacy aufweist. Es geht nicht ausschließlich um zu erwerbende Kompetenzen im Sinne des literacy-Konzepts. Vielmehr sollen die Schülerinnen und Schüler in der Auseinandersetzung mit den Versuchstationen der MINIPHÄNOMENTA erfahren, dass das Nachdenken über naturwissenschaftliche Fragestellungen spannend und interessant ist. Insofern kann die MINIPHÄNOMENTA auch als ein Beitrag zur Motivationsförderung und Interessenbildung in Bezug auf naturwissenschaftliche Denkweisen, Inhalte, Methoden und Fragestellungen verstanden werden. Im Rahmen der weiteren Argumentation konnte gezeigt werden, dass der Ansatz von Lück (2003, S. 18 ff.) eine geeignete bildungstheoretische Rahmenkonzeption für die MINIPHÄNOMENTA sein kann (vgl. Asmussen 2008).

Im Kontext dieses zweiten Aufsatzes zur MINIPHÄNOMENTA liegt der Fokus nicht mehr auf den konzeptionell-theoretischen Grundlagen. Vielmehr soll auf die empirische Begleitforschung zu dem Projekt eingegangen werden. Im Zentrum steht dabei die Frage der Wirksamkeit der MINIPHÄNOMENTA. Anders formuliert: Können Schülerinnen und Schüler der Primarstufe in einem derartig offenen Setting überhaupt etwas lernen? Negati-

ver: Besteht nicht die Möglichkeit, dass der Verzicht auf Versuchsanleitungen und Erklärungen an den Stationen dazu führt, dass die Schülerinnen und Schüler an ihnen ‚spielen‘, ohne dass dadurch ein wesentlicher Lerneffekt erzielt wird¹?

Dieser Fragestellung wird im Kontext dieses Artikels in drei Kapiteln nachgegangen: Zunächst wird auf den Forschungsstand eingegangen. In diesem Zusammenhang wird zum einen eine Skizze des empirischen Forschungsstandes innerhalb des Projekts geliefert, zum anderen werden Bezüge zu relevanten Diskursen hergestellt. Dieses Kapitel abschließend werden die beiden Fragestellungen der Untersuchung entwickelt. In dem folgenden methodischen Teil werden das Design, das Instrument, die Stichprobe, die Untersuchungsdurchführung und die Methoden der Datenanalyse beschrieben. Es folgt die Darstellung der Ergebnisse und abschließend deren Diskussion. Dabei wird ein besonderer Fokus auf Konsequenzen der Ergebnisse für die naturwissenschaftliche Grundbildung innerhalb des Sachunterrichts gelegt.

Forschungsstand und Theorie

Der Stand der empirischen Begleitforschung besteht wesentlich aus zwei in diesem Zusammenhang entstandenen Monographien:

Holst (2005) untersuchte in seiner Arbeit die Bildungswirksamkeit der MINIPHÄNOMENTA im Primar- und Orientierungsstufenbereich. Dabei wurden folgende interessante Ergebnisse erzielt: Die durchschnittliche Verweildauer der Kinder an den Versuchsstationen der MINIPHÄNOMENTA beträgt 1,3 Minuten. Dieser Wert mag zunächst relativ niedrig erscheinen. Hier muss jedoch der Kontext der Beobachtung berücksichtigt werden. Die Kinder hatten lediglich in den Pausenzeiten (häufig nur in den ‚großen‘ Pausen) Zugang zu den Versuchsstationen. Vergleicht man nun diesen Wert mit der durchschnittlichen Verweildauer von Schülerinnen und Schülern im Science-Center Phänomenta in Flensburg (unter einer Minute), so wird deutlich, dass es sich, so die Interpretation des Autors, um einen tendenziell eher hohen Wert handelt. Die Schülerinnen und Schüler arbeiteten an den Versuchsstationen nur selten allein. Etwa zwei Drittel der Kinder näherten sich den Stationen in Kleingruppen von drei bis fünf Mitgliedern. Sie diskutierten den Umgang mit den Versuchen. Oft stellten sich nach kurzer Zeit für das jeweilige Exponat Expertinnen und Experten heraus, deren Rat bei Bedarf von den Mitschülerinnen und Mitschülern eingeholt wurde. Die Kinder erarbeiteten in den Gruppen, oft im Prozess einer Diskussion, eigene Erklärungen für die jeweiligen Phänomene. Über die Verhaltensbeobachtung hinaus, verwendete Holst zur Analyse der Erklärungsmuster der Schülerinnen und Schüler die Methode der Concept-Maps (vgl. insbesondere Anderson-Inmann, Horny 1996). Die fachliche Bewertung der Erklärungen wurde auf einer Skala von 1 bis 3 (3 = beste fachliche Bewertung, 1 = schlechteste fachliche Bewertung) gemessen. Der Mittelwert aller Exponate lag bei 2,0. Dieses Ergebnis zeigt, dass sich die Erklärungsmuster der Schülerinnen und Schüler auf einem insgesamt angemessenen fachlichen Niveau befinden (zu einer genaueren Darstellung der Kriterien für die fachliche Bewertung der Erklärungen der Schülerinnen und Schüler vgl. Holst 2005).

Sauer (2005) analysierte den Einfluss der Versuchsstationen auf das naturwissenschaftlich-technische Lernen im Primarbereich. Der Autor führte insgesamt fünf Teiluntersuchungen auf der Basis von halbstandardisierten Interviews und Fragebögen durch. Dabei erscheinen insbesondere die folgenden Ergebnisse als interessant: 58% der Schülerinnen und Schüler gelang es nach dem Durchführen der Versuche, eine Station ihrer Wahl ihrem Aufbau und ihrer Funktion nach auch nach längerer Zeit (bis zu sechs Monate) korrekt zu beschreiben. 20% der Befragten konnten dies zumindest ansatzweise richtig tun. Es ließ sich aber nicht nur ein Einfluss auf die Schülerinnen und Schüler nachweisen. Auch der Unterricht veränderte sich durch die MINIPHÄNOMENTA: Inspiriert durch die Stationen veränderten die Lehrerinnen und Lehrer ihre Unterrichtsangebote. Naturwissenschaftliche Fragestellungen waren zum einen nach der MINIPHÄNOMENTA öfter Unterrichtsgegenstand. Zum anderen veränderte sich die Form des Unterrichts in Richtung eines forschend-entdeckenden Lernens (vgl. beispielsweise Ansari 2009).

Beide Studien weisen auf Wirkungen bzw. Einflüsse der MINIPHÄNOMENTA auf unterschiedliche Bereiche, insbesondere die Ebene der Schülerinnen und Schüler und die der Gestaltung des Unterrichts hin. Kritisch müssen in Bezug auf diese Ergebnisse jedoch zwei Aspekte angeführt werden: Zunächst kann festgestellt werden, dass die Untersuchungen vorrangig querschnittlich konzeptionalisiert sind. So steht eine Analyse der längerfristigen Wirksamkeit der MINIPHÄNOMENTA in weiten Teilen noch aus. Darüber hinaus spielen die methodisch-formalen Fähigkeiten der Kinder, das heißt ihre Fähigkeiten zu experimentieren und auf der Basis des Versuchs zu Erklärungen zu kommen, in beiden Untersuchungen eine nur untergeordnete Rolle (zu einer genaueren Beschreibung methodisch-formaler Fähigkeiten vgl. weiter unten in diesem Abschnitt). Betrachtet man jedoch die Konzeption der MINIPHÄNOMENTA, so liegt gerade hier eine zentrale Forschungsfrage. Die Schü-

¹ Es ist keinesfalls das Ziel des Autors die Begriffe Spielen und Lernen in einen Widerspruch zu bringen. Die obigen Ausführungen stellen vielmehr die im praktischen Einsatz der MINIPHÄNOMENTA geäußerten Befürchtungen von Eltern und teilweise auch Lehrerinnen und Lehrern dar.

lerinnen und Schüler erhalten im Kontext des Projekts keine fertigen Erklärungen, sondern sind aufgefordert, diese anhand der Auseinandersetzung mit den Versuchsstationen selbst zu entwickeln. In diesem Zusammenhang kommt der Frage, wie man zu diesen Erklärungsmustern kommt eine zentrale Bedeutung zu, denn die Exaktheit der Ergebnisse wird unter anderem wesentlich von den angewendeten Methoden und Denkweisen beeinflusst. Daher soll im Kontext dieser Arbeit der Frage nach den methodisch-formalen Fähigkeiten der Schülerinnen und Schüler vertiefend nachgegangen werden.

Die Nichtberücksichtigung methodenbezogener Fragen in den genannten Untersuchungen erscheint umso erstaunlicher, da eben diese Fragen in einer Vielzahl aktueller bildungswissenschaftlicher und bildungswissenschaftlich relevanter Diskurse in ganz unterschiedlicher Form eine zentrale Rolle spielen. Zur Skizzierung seien an dieser Stelle einige Beispiele aus dem Elementar- und Primarbereich angeführt:

- 1) Psychologische Grundlagen methodischen, insbesondere naturwissenschaftlichen Denkens: Im Zusammenhang dieses Diskurses (zu einem Überblick vgl. Sodian, Koerber, Thoermer 2006) wird aus der Perspektive der Kognitionspsychologie der Frage nachgegangen, inwieweit Kinder in der Lage sind grundsätzliche naturwissenschaftliche Denkweisen, wie zum Beispiel Prinzipien der Variablenkontrolle zu verstehen und anzuwenden.
- 2) Stärkung der lernmethodischen Kompetenz im Elementarbereich: Basierend auf der unauflösbaren Verbindung von inhaltlichem und methodischem Lernen schlagen Samuelsson und Pramling (2007) aus der Sicht der Erziehungswissenschaft ein metakognitives Modell vor, das inhaltliches und methodisches Lernen verbindet. Die Kernaussage lässt sich dabei wie folgt zusammenfassen: „Die Kinder reflektieren gemeinsam mit der Hilfe der Erzieherin, das sie lernen, was sie lernen und wie sie lernen“ (Gisbert 2003, S. 93).
- 3) Vermittlung eines Wissenschaftsverständnisses in Bezug auf die Naturwissenschaften in der Primarstufe (engl. nature of science): Unter dem Terminus „Wissenschaftsverständnis“ verstehen Sodian, Thoermer, Kircher, Grygier und Günther (2002, S. 192) aus der Perspektive der Didaktik der Naturwissenschaften eine Einsicht in erkenntnistheoretische, wissenschaftstheoretische und wissenschaftsethische Grundlagen der Naturwissenschaften.

Insbesondere der letztgenannte Punkt ist für den vorliegenden Untersuchungszusammenhang von zentraler Bedeutung. Im Rahmen unterschiedlicher vorrangig qualitativer Interventionsstudien (vgl. Sodian, Thoermer, Kircher, Grygier, Günther 2002, S. 192; ebd., S. 59 ff.) konnte gezeigt werden, dass die Anbahnung eines Wissenschaftsverständnisses in der Primarstufe möglich ist.

Daher scheint eine Untersuchung des Einflusses der MINIPHÄNOMENTA auf die methodisch-formalen Fähigkeiten der Schülerinnen und Schüler (vgl. unten in diesem Abschnitt) vorstellbar. Die Perspektive muss hier jedoch erheblich eingeschränkt werden. Erkenntnistheoretische und wissenschaftsethische Fragestellungen spielen im Kontext der MINIPHÄNOMENTA keine Rolle. Zentral ist aber der wissenschaftstheoretische Aspekt unter der Fragestellung „Wie experimentieren die Kinder?“ Genauer: Können sie in der Auseinandersetzung mit den Stationen der MINIPHÄNOMENTA methodisch-formale Fähigkeiten aufbauen? Darüber hinaus soll untersucht werden, ob die Wirkung der MINIPHÄNOMENTA auf die methodisch-formalen Fähigkeiten der Schülerinnen und Schüler über einen längeren Zeitraum erhalten bleibt.

Die Variable methodisch-formale Fähigkeiten wird im Rahmen dieser Untersuchung vom Autor als ein komplexes Merkmal operational definiert, dass durch die folgenden Subskalen gemessen werden kann:

- 1) Logische Fähigkeiten
- 2) Fähigkeiten des spezifisch-naturwissenschaftlichen Denkens
- 3) Experimentelle Fähigkeiten

Diese drei Subskalen werden im Folgenden näher erläutert. Der erste Teilbereich, die logischen Fähigkeiten kennzeichnen sich wie folgt: „Logik ist die Lehre vom richtigen Denken, genauer von den Formen und Methoden (also nicht dem Inhalt), des richtigen Denkens. Sie kann nicht zeigen, was man Denken muss, sondern nur, wie man von irgendeinem Gegenstand ausgehend, denkend fortschreiten muss, um zu richtigen Ergebnissen zu gelangen“ (Störing 1987, S. 177). Von besonderer Bedeutung für diesen Untersuchungszusammenhang ist die klassische Unterscheidung zwischen dem induktiven und dem deduktiven Schluss (vgl. Speck, Acham 1980, Chalmers 2001). Es steht in dieser Subskala also die Frage im Vordergrund wie Kinder zu Ergebnissen kommen. Folgen sie dabei korrekt den Regeln des Schließens oder ergeben sich logische Indifferenzen. Ein Blick in die entwicklungspsychologische Forschung zu diesem Bereich ergibt ein heterogenes Bild. Der bisherige Forschungsstand zum deduktiven Denken von Kindern ist äußerst komplex und vielfältig. Es spricht aber dennoch vieles dafür, dass Kinder ab dem Alter von acht Jahren einfache deduktive Schlüsse ziehen können (vgl. Oerter, Montada 1995, S. 593 ff.). In Bezug auf das induktive Schließen ergibt sich ein anderes Bild. Kinder scheinen, die diesem Schluss zugrunde liegende Fähigkeit des Vergleiches, schon deutlich früher zu entwickeln (vgl. Oerter, Montada 1995, S. 599 ff.).

Die zweite in dieser Variablen erfasste Subskala betrifft die spezifisch-naturwissenschaftlichen Fähigkeiten. Sie lässt sich beispielhaft durch folgende Fragestellungen kennzeichnen: Gelingt den Kindern typisch-naturwissenschaftliches Denken, nehmen sie also zum Beispiel eine Verknüpfung von Empirie und Theorie vor,

erkennen und trennen sie Variablen, verstehen sie das Prinzip der kausalen Verursachung? Zu diesem Teilbereich hat sich gerade in den letzten Jahren, insbesondere im Zusammenhang der Münchner Längsschnittstudie LOGIK (zu einem zusammenfassenden Überblick mit Fokus auf die naturwissenschaftliche Grundbildung vgl. Sodian, Koerber, Thoermer 2006) ein breiter Forschungsbestand entwickelt. Kinder im Alter der Primarstufe zeigen beispielsweise, so ein zentrales Ergebnis, in assistierten Aufgaben-Kontexten ein Grundverständnis des Vorgangs der Variablenprüfung. Auch gelingt ihnen in Ansätzen eine Theorie-Evidenz-Verknüpfung.

Die letzte Subskala beschreibt die experimentellen Fähigkeiten. Besitzen die Kinder experimentelle Fähigkeiten, manipulieren sie also die Variablen, gehen sie systematisch vor, benötigen sie wenig Unterstützung, arbeiten sie sorgfältig usw.? Diese dritte Subskala weist inhaltliche Schnittmengen mit der ersten und zweiten auf, betont aber eher den „handwerklich-praktischen Anteil“ des Experimentierens.

Auf der Basis des Erkenntnisinteresses dieses Aufsatzes und der vorgestellten Definition der Variable methodisch-formale Fähigkeiten ergeben sich die folgenden Fragestellungen der Untersuchung:

- 1) Verändern sich die methodisch-formalen Fähigkeiten der Schülerinnen und Schülern durch die Teilnahme an der MINIPHÄNOMENTA?
- 2) Gibt es eine Abhängigkeit der Wirksamkeit der MINIPHÄNOMENTA von der Zeit?

Methode

Untersuchungsdesign

Das Untersuchungsdesign zur Analyse der beiden Fragestellungen ist zweistufig angelegt:

Untersuchungsstufe I: Hier handelt es sich nach der Definition von Bortz und Döring (2006) um einen quasi-experimentellen Zwei-Gruppen-Prätest-Posttest-Plan mit zwei Messwiederholungen:

		t ₋₁ 1 Monat	t ₀	t ₊₃ 3 Monate	t ₊₆ 6 Monate
3. Klasse	Experimentalgruppe I (3)	Prätest	Treatment	Posttest I	Posttest II
	Kontrollgruppe (3)	Prätest			Posttest II'
4. Klasse	Experimentalgruppe I (4)	Prätest	Treatment	Posttest I	Posttest II
	Kontrollgruppe (4)	Prätest			Posttest II'

Tabelle 1: Das Design der Untersuchungsstufe I

Die Untersuchung findet in der dritten und vierten Klassenstufe statt. Für beide Klassenstufen werden jeweils eine Experimental- und eine Kontrollgruppe gebildet. Die Zuweisung der Teilnehmerinnen und Teilnehmer geschieht durch ein Kriterium: Die Teilnahme beziehungsweise Nichtteilnahme an der MINIPHÄNOMENTA. Es sind drei Messzeitpunkte für die Experimentalgruppen (zwei davon auch für die Kontrollgruppen) vorgesehen:

- 1) Prätest: Ein Monat vor der Teilnahme der Experimentalgruppe I an der MINIPHÄNOMENTA
- 2) Posttest I: Drei Monate nach der Teilnahme an der MINIPHÄNOMENTA
- 3) Posttest II: Sechs Monate nach der Teilnahme an der MINIPHÄNOMENTA

Mit dem dargestellten Design lassen sich die Wirkungen der MINIPHÄNOMENTA über einen Zeitraum von insgesamt sechs Monaten messen.

Untersuchungsstufe II: Bei dieser Untersuchungsstufe handelt es sich um einen Ein-Gruppen-Prätest-Posttestplan (vgl. Bortz, Döring 2006), der im Rahmen der Untersuchung eine explorative Funktion erfüllt:

		Prätest	t ₀	t _{+2 1/4} Jahre (Posttest)
3. Klasse	Experimentalgruppe II (3)	Prätestergebnisse der <i>Untersuchungsstufe I</i> <i>siehe Tabelle 1</i>	Treatment	Posttest
4. Klasse	Experimentalgruppe II (4)	Prätestergebnisse der <i>Untersuchungsstufe I</i> <i>siehe Tabelle 1</i>	Treatment	Posttest

Tabelle 2: Das Design der Untersuchungsstufe II

Das Design wechselt an dieser Stelle von einem längsschnittlichen in ein querschnittliches Vorgehen. Prä- und Posttest der Untersuchungsstufe II werden also nicht mehr an der gleichen Gruppe, sondern an zwei unterschiedlichen Gruppen erhoben. Als Prätestergebnisse der Untersuchungsstufe II werden die Prätestergebnisse der Untersuchungsstufe I verwendet. Diese werden mit den Ergebnissen einer zweiten Gruppe (Posttest), zusammengesetzt aus Schülerinnen und Schülern der dritten und vierten Klasse verglichen, die vor 27 Monaten an der MINIPHÄNOMENTA teilgenommen hat. So kann ein *erster* Überblick über die Wirksamkeit der MINIPHÄNOMENTA über einen Zeitraum von zwei einviertel Jahren gewonnen werden.

Untersuchungsinstrument

Auf der Basis der drei skizzierten Subskalen (logische Fähigkeiten, Fähigkeiten des spezifisch-naturwissenschaftlichen Denkens und experimentelle Fähigkeiten) wurde ein Untersuchungsinstrument sowohl für den Prä- als auch für den Posttest entwickelt, das Beobachtungs- und Interviewdaten erfasst.

Im Rahmen der Verhaltensbeobachtung werden die Kinder in ihrem Umgang mit dem Versuch beobachtet. In dem Interviewteil sollen die Kinder zum Beginn des Versuches Hypothesen äußern und zum Ende der Untersuchung ihre bisherigen Denkbemühungen zusammenfassend verbalisieren sowie sich mit Fragen zu wesentlichen Grundlagen naturwissenschaftlichen Denkens auseinandersetzen.

Das Untersuchungsinstrument besteht aus 15 gleichberechtigten, nicht hierarchisierten Kategorien. Sie wurden im Vorfeld der Untersuchung in Zusammenarbeit mit Lehrerinnen und Lehrern des Fachs Sachunterricht in Schleswig-Holstein entwickelt. Die folgende Tabelle zeigt einige Beispielkategorien aus dem Untersuchungsinstrument:

	logische Fähigkeiten	Fähigkeiten des spezifisch-naturwissenschaftlichen Denkens	experimentelle Fähigkeiten
Beispiel-Kategorie 1	Die Versuchsperson kann deduktiv schließen.	Die Versuchsperson lässt Ansätze der Variablentrennung erkennen.	Die drei Variablen werden manipuliert ² .
Beispiel-Kategorie 2	Die Ergebnisse werden korrekt verknüpft.	Die Versuchsperson beobachtet genau.	Die Manipulation der Variablen erfolgt mit der notwendigen Gründlichkeit.
Beispiel-Kategorie 3	Die Repräsentation der Variablen gelingt.	Das Prinzip der Kausalität wird verstanden.	Die Möglichkeiten des empirischen Materials werden ausgeschöpft.

Tabelle 3: Beispiel-Kategorien

Die Zuweisung der Personendaten zu den Likert-skalierten Kategorien durch die Versuchsleiterin / den Versuchsleiter geschieht auf der Basis eines Zuweisungsschemas. Sie wird in der Situation von der Versuchsleiterin / dem Versuchsleiter vorgenommen. Die Tabelle vier zeigt das Zuweisungsschema für die erste, in Tabelle drei genannte Beispielkategorie in den drei Subskalen:

² Bei jedem, der zu den unterschiedlichen Untersuchungszeitpunkten verwendeten Versuche können drei Variablen leicht erkannt und manipuliert werden. Zu einer genaueren Beschreibung der verwendeten Versuche sowie dem genauen Untersuchungsablauf vgl. den Abschnitt „Untersuchungsdurchführung“.

	4 Punkte	3 Punkte	2 Punkte	1 Punkt
Die Versuchsperson kann deduktiv schließen.	Mindestend zwei Hypothesen werden abgeleitet und geprüft.	Mindestens eine Hypothese wird abgeleitet und geprüft.	Mindestens eine Hypothese wird teilweise korrekt abgeleitet und überprüft.	Deduktives Schließen gelingt nicht oder nur sehr rudimentär.
Die Versuchsperson lässt Ansätze der Variablentrennung erkennen.	Das Prinzip der Variablentrennung wird durchgängig angewandt.	Die Variablentrennung wird in mindestens zwei Fällen durchgeführt.	Die Variablen werden in einem Fall getrennt.	Es gibt keine Variablentrennung.
Die drei Variablen werden manipuliert.	Die Schülerinnen und Schüler manipulieren alle drei Variablen zielgerichtet.	Die Schülerinnen und Schüler manipulieren drei Variablen.	Die Schülerinnen und Schüler manipulieren zwei Variablen.	Die Schülerinnen und Schüler manipulieren weniger als zwei Variablen.

Tabelle 4: Beispiele für das Zuweisungsschema

Um eine möglichst hohe Übereinstimmung zwischen den einzelnen Beobachterinnen / Beobachtern zu erreichen, wurde im Vorfeld der Untersuchung eine forschungsmethodische Schulung für die drei beteiligten Untersuchungsleiterinnen / Untersuchungsleiter durchgeführt. Neben einer kurzen Einführung in den Untersuchungszusammenhang stand hier das gemeinsame Zuweisen von Personendaten zu den Kategorien auf der Basis der Kodieranweisung im Vordergrund. Dies geschah anhand mehrerer Videobeispiele.

Zusätzlich zur Verhaltensbeobachtung wurde ein Interview durchgeführt. Dieses besteht aus drei Abschnitten: In einem ersten Schritt wurden die Schülerinnen und Schüler gebeten Hypothesen über die im Experiment ersichtlichen Zusammenhänge zu bilden. Nach der Beendigung ihrer experimentellen Arbeit erhielten die Schülerinnen und Schüler die Aufgabe ihren Arbeitsprozess zusammenzufassen und ihre wesentlichen experimentellen Schritte zu begründen. In einem dritten Schritt wurden die Schülerinnen und Schüler zu den Grundzügen des naturwissenschaftlichen Experimentierens befragt. In diesem Zusammenhang sollten sie sich mit folgenden Fragen auseinandersetzen: (1) Was ist ein naturwissenschaftliches Experiment? und (2) Warum macht man naturwissenschaftliche Experimente?

Die Aussagen der Schülerinnen und Schüler wurden aufgenommen, das heißt die Zuweisung geschah, abweichend von der Verhaltensbeobachtung, nicht in der Interviewsituation, sondern wurde später vorgenommen. Die im Zusammenhang des Interviews gewonnenen Daten werden in das Untersuchungsinstrument integriert. Dabei gilt für die Interviewschritte ein unterschiedliches Vorgehen:

- 1) Zusammenfassung der bisherigen Denkbemühungen: Hier gilt der Grundsatz, dass im Falle von Abweichungen zwischen dem vorliegenden Beobachtungsprotokoll und der Tonbandaufnahme in den einzelnen Kategorien der Punktwert um einen Testpunkt in jede Richtung verändert werden konnte.
- 2) Hypothesenbildung und Grundzüge des naturwissenschaftlichen Experimentierens: Hierzu liegen aus der Verhaltensbeobachtung keine Daten vor. Auch für diese Kategorien wurden die Aussagen der Schülerinnen und Schüler anhand eines Zuweisungsschemas bewertet und in das Kategoriensystem integriert.

Die folgende Skizze stellt den strukturellen Aufbau und das zu Stande kommen des Rohpunktwertes als eine Integration aus Beobachtungs- und Interviewdaten nochmals grafisch dar:

L A B O R P S Y C H O L O G I S C H E S S E T T I N G	Verhaltensbeobachtung der Kinder beim Experimentieren	Bewertung in Kategorien
	Interviewteile:	Bewertung in Kategorien
	- Aufforderung zur Hypothesenbildung	Bewertung in Kategorien/ Vergleich mit Beobach- tungsdaten
	- Zusammenfassung des Arbeitsprozesses	Bewertung in Kategorien
	- Fragen zu den Grundzügen des naturwissenschaftlichen Experimentierens	Bewertung in Kategorien
		Summe: Punktwert

Grafik 1: Das Untersuchungsinstrument

Das Untersuchungsinstrument wurde einem Prätest unterzogen. Im Vordergrund stand dabei zum einen die Untersuchung der Praktikabilität des Verfahrens, zum anderen die Berechnung zentraler Kennwerte: Itemparameter (Trennschärfe und Schwierigkeit) und das Testgütekriterium der Reliabilität.

Die Schwierigkeit einer Kategorie gibt an von wie vielen Versuchspersonen eine Kategorie gelöst bzw. nicht gelöst wurde. Beträgt die Schwierigkeit eins wurde die Kategorie von allen Versuchspersonen gelöst. Im Falle dieser Untersuchung bedeutet dies: Alle Versuchspersonen haben in der Kategorie die volle Punktzahl erreicht. Um eine ausreichende Differenzierungsfähigkeit der Kategorien zu erreichen, sollte die Schwierigkeit zwischen 0,3 und 0,8 liegen. Die durchschnittliche Schwierigkeit der Kategorien in der vorliegenden Untersuchung beträgt 0,5, wobei die Schwierigkeit der einzelnen Kategorien stets zwischen 0,3 und 0,8 liegt. Damit ist eine ausreichende Differenzierungsfähigkeit der Kategorien gewährleistet. In einem nächsten Schritt wurde die Trennschärfe der Kategorien ermittelt. Sie basiert auf einer Produkt-Moment-Korrelation zwischen dem erreichten Punktwert in der Kategorie und dem Gesamtpunktwert. Sie gibt also an, wie gut die einzelne Kategorie den Gesamtpunktwert repräsentiert. Mit einer durchschnittlichen Trennschärfe von 0,7 ist dieser Wert ebenfalls befriedigend. Die Reliabilität des Verfahrens wurde auf der Basis des Cronbachs-Alpha-Index ermittelt. Dieser liegt mit einem Wert von $\alpha=0,9$ hoch.

Stichprobenkonstruktion

Bei der Stichprobe handelt es sich um eine dreistufig gezogene Zufalls-Cluster-Stichprobe.

Insgesamt nahmen in der Untersuchungsstufe I 155 Schülerinnen und Schüler ($AM_{\text{Alter}} = 9,1$ Jahre) der dritten und vierten Klassenstufe an der Untersuchung teil. Dabei entfielen 75 Schülerinnen und Schüler auf die Experimentalgruppe (Klasse 3 = 36, Klasse 4 = 39). 80 Schülerinnen und Schüler sind der Kontrollgruppe zuzurechnen (Klasse 3 = 38, Klasse 4 = 42). An der Untersuchungsstufe II nahmen in der Experimentalgruppe 28 Schülerinnen und Schüler ($AM_{\text{Alter}} = 8,9$ Jahre) teil. Dabei waren 15 Schülerinnen und Schüler in der dritten Klasse. 13 Schülerinnen und Schüler waren in der vierten Klasse.

Untersuchungsdurchführung

Grundlage der Untersuchung ist die Bearbeitung einer experimentellen Aufgabe durch die Schülerinnen und Schülern in einem laborpsychologischen Rahmen. Zur Minimierung von Testungseffekten werden bei den drei Untersuchungsterminen der Untersuchungsstufe I unterschiedliche experimentelle Aufgaben eingesetzt. Zur Sicherstellung der Vergleichbarkeit der Versuche wurden für deren Auswahl die folgenden Kriterien angelegt:

- 1) Es muss eine konkrete, für Kinder fassbare Aufgabe gelöst werden.
- 2) Drei, klar unterscheidbare Variablen sind manipulierbar.

Zur Verdeutlichung sei die Prätest-Aufgabe, das „Vario-Pendel“³ angeführt: Es handelt sich hierbei um einen Aufbau, der aus zwei Pendeln besteht. Die Aufgabe besteht darin, beide Pendel in Phase schwingen zu lassen. Dazu können verändert werden: Die Fadenlänge, das Gewicht der Pendelkörper und der Anfangspunkt der Pendelauslenkung.

Die formalen Fähigkeiten der Schülerinnen und Schüler werden durch die Kombination einer Verhaltensbeobachtung mit einem Interview im Kontext der experimentellen Aufgabe erhoben (vgl. oben). Die folgende Tabelle schildert den Ablauf der Untersuchung:

	Programmpunkt	Inhalt
1	Begrüßung	Die Schülerinnen und Schüler werden am Versuchsaufbau begrüßt.
2	Aufgabenstellung mit Demonstration	Zunächst wird den Schülerinnen und Schülern die Aufgabe vorgestellt. Es folgt die Nennung und die Demonstration der drei zu manipulierenden Variablen durch die Versuchsleiterin / den Versuchsleiter.
3	Aufforderung zur Hypothesenbildung	Die Schülerinnen und Schüler werden aufgefordert Vermutungen zu äußern und zu begründen, welche Variable verändert werden muss, um die Pendel mit gleicher Periodendauer schwingen zu lassen.
4	Schülerinnen und Schüler experimentieren	Die Schülerinnen und Schüler experimentieren bis zu dem Punkt, an dem sie mit der Lösung der Aufgabe zufrieden sind.
5	Ergebniszusammenfassung	Die Schülerinnen und Schüler werden gebeten zusammenzufassen, was sie getan haben und warum sie es getan haben. Es folgt ein Fazit, wie die Aufgabe gelöst werden konnte.
6	weitere Fragen	In diesem Zusammenhang wird zunächst auf offene Fragen und Widersprüche eingegangen. Darauf aufbauend werden Fragen zum spezifisch naturwissenschaftlichen Denken gestellt. Die Schülerinnen und Schüler werden gebeten, sich Gedanken zu machen, was ein Experiment ist und warum Naturwissenschaftlerinnen und Naturwissenschaftler Experimente machen.

Tabelle 5: Ablauf der Datenerhebung

Datenaufbereitung

Die Untersuchungsprotokolle, die eine Kombination aus Beobachtungs- und Interviewdaten enthalten, wurden in das Statistikprogramm SPSS eingegeben. Die Datenanalyse erfolgte auf der Basis der sich so ergebenden Matrix.

Datenanalyse

Die Überprüfung der beiden Fragestellungen erfolgt methodisch als Analyse eines Prä-Posttest-Unterschiedes. Die Untersuchung gliedert sich dabei in die folgende Sequenz:

- 1) Ermittlung des Prä-Posttest-Unterschiedes beziehungsweise des Nettoeffekts zur Deskription der Daten (zu dem Begriff des Nettoeffekts vgl. unten)
- 2) Signifikanzüberprüfung
- 3) Bestimmung der Effektgröße
- 4) Umfang der Wirkung des Treatments: In diesem Zusammenhang wird der Frage nachgegangen, ob die Wirksamkeit der MINIPHÄNOMENTA auf einzelne Subskalen der abhängigen Variablen beschränkt war oder ob sich die Wirkung über alle Subskalen erstreckt.

Der Terminologie und Logik des sozialwissenschaftlichen Experiments folgend, kann zwischen zwei unterschiedlich methodisch kontrollierten Ebenen der Wirksamkeit unterschieden werden:

- 1) Ebene I: Wirksamkeitsüberprüfung als Haupteffekt
- 2) Ebene II: Wirksamkeitsüberprüfung als Interaktionseffekt

Ad 1: *Wirksamkeitsüberprüfung als Haupteffekt*: Wirksamkeit ist hier definiert als die Veränderung eines Merkmales nach der Zeit. Es wird folglich von der Wirksamkeit der MINIPHÄNOMENTA ausgegangen, sofern eine signifikante Prä-Posttest-Differenz mit einer ausreichenden Effektstärke (zur Kategorisierung der Effekt-

³ Das Exponat wurde mir für den Zeitraum der Untersuchung freundlicherweise von der Phänomenta (Norderstr. 157 – 163, 24943 Flensburg) durch Herrn Achim Englert zur Verfügung gestellt.

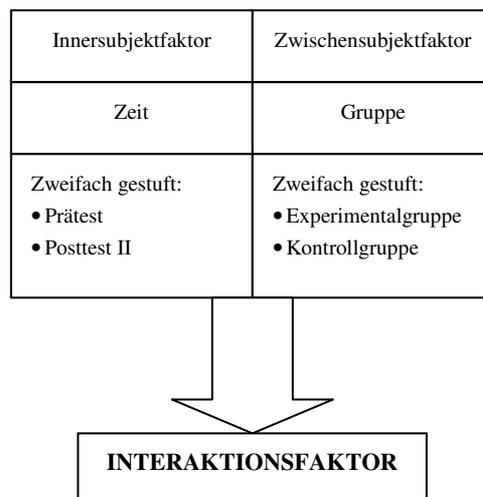
stärken vgl. unten) vorliegt. Diese Wirksamkeitsebene wird aufgrund der methodischen Beschaffenheit der Studie lediglich für die Differenz Prätest / Posttest I (Untersuchungsstufe I) und die Prä-Posttest-Differenz der Untersuchungsstufe II ermittelt, da in beiden Fällen keine Ergebnisse in der Kontrollgruppe erhoben werden. Zur Überprüfung auf statistische Bedeutsamkeit wird ein t-Test für unabhängige Stichproben verwendet. Die Effektgröße wird mit Cohens d-Maß berechnet. Die Darstellung erfolgt als Überlappungsbereich der beiden Verteilungen (beispielsweise Prä- und Posttest). Dabei ist zu beachten, dass folgende Beziehung gilt: Je kleiner der Überlappungsbereich, desto größer ist die Wahrscheinlichkeit, dass es sich um zwei unterschiedliche Verteilungen handelt. In der Folge bedeutet dies, dass ein kleiner Überlappungsbereich einen großen Effekt bedeutet.

Ad 2: *Wirksamkeitsüberprüfung als Interaktionseffekt*: Zum Nachweis dieser Wirksamkeitsebene werden über zwei Vergleichszeitpunkte hinaus die Daten einer Kontroll- und einer Experimentalgruppe benötigt. Untersuchungsbedingt kann diese Methode folglich nur in Bezug auf die Differenz zwischen dem Prätest und dem Posttest II eingesetzt werden. Zum Nachweis eben dieser Wirksamkeitsebene wird als deskriptiver Wert zunächst der Nettoeffekt nach Bortz und Döring (2006) berechnet. Diesem Vorgehen liegt das folgende Berechnungsschema zugrunde:

	Prätest	Posttest I	Posttest II	
Experimentalgruppe	arithmetisches Mittel I	hier irrelevant	arithmetisches Mittel III	E = arithmetisches Mittel I - arithmetisches Mittel III
Kontrollgruppe	arithmetisches Mittel I'	nicht erhoben	arithmetisches Mittel III'	K = arithmetisches Mittel I' - arithmetisches Mittel III'
Nettoeffekt:				E - K

Tabelle 6: Bestimmung des Netto-Effekts

Die zugrunde liegende Unterschiedshypothese kann aufgrund des Vorhandenseins von zwei unabhängigen Variablen (Zeit und Gruppe) nicht mehr mit einem t-Test überprüft werden. Zur Anwendung kommt vielmehr eine zweifaktorielle Varianzanalyse. Mit Hilfe dieses Verfahrens können Aussagen darüber gemacht werden, welcher Anteil der Varianz durch den Innersubjektfaktor (Zeit: Prätest, Posttest II), den Zwischensubjektfaktor (Gruppe: Experimentalgruppe, Kontrollgruppe) und die Interaktion der beiden Faktoren hervorgerufen wird. Die beiden Haupteffekte können Bezug nehmend auf die Definition der Ebene II der Wirksamkeit (vgl. Tabelle 8) ignoriert werden. Von Interesse ist daher nur der Interaktionseffekt. Die folgende Grafik stellt diesen Zusammenhang nochmals dar:



Der Interaktionsfaktor wird auf Basis des Allgemeinen Linearen Modells bestimmt. Aufgrund seiner Robustheit wurde als Prüfgröße die Pilai-Spur ausgewählt.

Damit stehen die Elemente zur Überprüfung der Fragestellungen bereit. Es gilt nun die Stufen der Wirksamkeit für beide Ebenen operational zu definieren. Dies geschieht anhand der beiden folgenden Tabellen:

Ebene I: Wirksamkeit als Haupteffekt

Wirksamkeitsstufe	Bezeichnung	Definition
Wirksamkeitsstufe 1	UNWIRKSAM	Die Unterschiede zwischen Prä- und Posttest sind nicht signifikant.
Wirksamkeitsstufe 2	TEILWEISE WIRKSAM	Mindestens signifikanter Unterschied Effektstärke: Cohens d-Maß dargestellt als Überlappungsbereich; nach Bortz und Döring (2006, S. 607 f.): 69% bis 88% ⁴ Umfang der Wirkung: Wirksamkeit auf einer Subskala
Wirksamkeitsstufe 3	WIRKSAM	Mindestens signifikanter Unterschied Effektstärke: Cohens d-Maß dargestellt als Überlappungsbereich; nach Bortz und Döring (2006, S. 607 f.): 69% bis 88% Umfang der Wirkung: Wirksamkeit auf zwei Subskalen
Wirksamkeitsstufe 4	HOCHWIRKSAM	Mindestens signifikanter Unterschied Effektstärke: Cohens d-Maß dargestellt als Überlappungsbereich; nach Bortz und Döring (2006, S. 607 f.): 0% bis 68% ⁵ Umfang der Wirkung: Wirksamkeit auf drei Subskalen

Tabelle 7: Wirksamkeitsstufen der Ebene I

Ebene II: Wirksamkeit als Interaktionseffekt

Wirksamkeitsstufe	Bezeichnung	Definition
Wirksamkeitsstufe 1'	UNWIRKSAM'	Der Interaktionseffekt ist nicht signifikant
Wirksamkeitsstufe 2'	TEILWEISE WIRKSAM'	Mindestens signifikanter Interaktionseffekt Effektstärke: $\eta^2_{\text{part.}}$ nach Bortz und Döring (2006, S. 606): 0,10 bis 0,24 Umfang der Wirkung: Wirksamkeit auf einer Subskala
Wirksamkeitsstufe 3'	WIRKSAM'	Mindestens signifikanter Interaktionseffekt Effektstärke: $\eta^2_{\text{part.}}$ nach Bortz und Döring (2006, S. 606): 0,10 bis 0,24 Umfang der Wirkung: Wirksamkeit auf zwei Subskalen
Wirksamkeitsstufe 4'	HOCHWIRKSAM'	Mindestens signifikanter Interaktionseffekt Effektstärke: $\eta^2_{\text{part.}}$ nach Bortz und Döring (2006, S. 606): größer als 0,24 Umfang der Wirkung: Wirksamkeit auf drei Subskalen

Tabelle 8: Wirksamkeitsstufen der Ebene II

⁴ Bei der Interpretation der Werte ist die antiproportionale Beziehung zwischen dem Ausmaß der Überlappung der Verteilungen und der Höhe des Effekts zu beachten. Es gilt: Je kleiner der Überlappungsbereich, desto größer die Effektgröße und folglich desto größer der Einfluss des treatments MINIPHÄNOMENTA.

⁵ Bei dem Wert 0% Überlappung handelt es sich um einen theoretischen Wert. Basierend auf der Annahme des zentralen Grenzwerttheorems handelt es sich bei n größer 30 um Normalverteilungen. Diese nähern sich nur der X-Achse, schneiden diese aber nie. In der Folge können sich nur sehr kleine Werte aber niemals Werte von 0% ergeben. Trotz dieser theoretischen Einschränkung wird hier aus Gründen der Praktikabilität der Wert 0% gewählt.

Ergebnisse

Die folgende Tabelle stellt einleitend die arithmetischen Mittel der Kontroll- und Experimentalgruppe zu den einzelnen Untersuchungszeitpunkten dar:

	Untersuchungsstufe I			Untersuchungsstufe II
	Prätest	Posttest I	Posttest II	Posttest
Arithmetisches Mittel - Experimentalgruppe	33,4 Punkte	50,5 Punkte	49,3 Punkte	46,7 Punkte
Arithmetisches Mittel - Kontrollgruppe	33,3 Punkte	---	35,7 Punkte	---

Tabelle 9: Deskriptive Werte

Die Schülerinnen und Schüler erreichten im Prätest mit 33,4 Punkten (Experimentalgruppe) beziehungsweise 33,3 Punkten (Kontrollgruppe) etwas mehr als die Hälfte der insgesamt verfügbaren Punkte. Diese sehr ähnlichen Punktwerte im Vorfeld der Untersuchung unterstreichen im Hinblick auf die Stichprobenkonstruktion die Vergleichbarkeit der beiden Gruppen in Bezug auf das Merkmal methodisch-formale Fähigkeiten.

Die erzielten Punktwerte in der Experimentalgruppe änderten sich im weiteren Verlauf deutlich. Im Posttest I, Posttest II und in der Untersuchungsstufe II erzielten die Schülerinnen und Schüler bessere Ergebnisse, auch wenn die Punktwerte mit dem zeitlichen Abstand zum treatment wieder leicht abfallen. In der Kontrollgruppe ist ein solcher Punktanstieg nicht zu verzeichnen. Die Schülerinnen und Schüler erreichten im Posttest II in Relation zum Prätest nur leicht bessere Werte, was vermutlich auf Testwiederholungseffekte zurück zu führen sein dürfte. Die Verbesserung der methodisch-formalen Fähigkeiten in der Experimentalgruppe beim Ausbleiben derartiger Effekte in der Kontrollgruppe legt die Wirksamkeit des treatments auf die abhängige Variable nahe. Im Einzelnen ergeben sich die folgenden Differenzen:

Untersuchungsstufe I

Prätest - Posttest I: 17,1 Punkte

Prätest - Posttest II (Nettoeffekt – vgl. Tabelle 6): -13,5 Punkte

Untersuchungsstufe II

Prätest - Posttest: 13,3 Punkte

Im weiteren Textverlauf soll nun mit der inferenzstatistischen Auswertung fortgefahren werden. Die folgende Tabelle fasst die Ergebnisse der Signifikanzüberprüfung und die ermittelten Effektstärken zusammen:

	Untersuchungsstufe I		Untersuchungsstufe II
	Prätest - Posttest I	Prätest - Posttest II	Prätest - Posttest
Signifikanzniveau	hoch signifikant $p \leq 0,01$	hoch signifikant $p \leq 0,01$	hoch signifikant $p \leq 0,01$
Effektstärke	Überlappungsbereich: 42%	$\eta^2_{\text{part.}} = 0,32$	Überlappungsbereich: 47%

Tabelle 10: Inferenzstatistische Auswertung⁶

Die in der Tabelle dargestellten Ergebnisse unterstreichen das bereits in der Deskription der Werte angedeutete Ergebnis: Die Wirksamkeit des treatments MINIPHÄNOMENTA auf die methodisch-formalen Fähigkeiten der Schülerinnen und Schüler. So liegen in allen drei Fällen hoch signifikante Unterschiede zum Prätest mit hohen Effektstärken vor. Von besonderer Bedeutung ist in diesem Zusammenhang das Ergebnis des Prätests II der Untersuchungsstufe I, da die hier erzielten Ergebnisse zusätzlich durch eine Kontrollgruppe abgesichert sind. Durch dieses vollständige Experiment, hier Wirksamkeit als Interaktionseffekt genannt, wird der Einfluss des treatments auf einer methodologisch abgesicherteren Ebene verdeutlicht.

Nach der Herausarbeitung dieser Datentendenzen soll nun auf die Ergebnisse in den einzelnen Subskalen (logische Fähigkeiten, spezifisch-naturwissenschaftliche Fähigkeiten, experimentelle Fähigkeiten) eingegangen werden. Die folgende Tabelle stellt die Ergebnisse zusammenfassend dar:

⁶ Zu den angewendeten statistischen Verfahren vgl. Abschnitt „Datenanalyse“

	Untersuchungsstufe I			Untersuchungsstufe II ⁷
	Prätest ⁸	Posttest I	Posttest II (Nettoeffekt)	Posttest
Logische Fähigkeiten	12,0 Punkte	17,5 Punkte	-5,6	---
spezifisch-naturwissenschaftliches Denken	6,6 Punkte	11,0 Punkte	-4,5	---
experimentelle Fähigkeiten	14,8 Punkte	21,3 Punkte	-4,6	---

Tabelle 11: Weitere Befunde zu den Subskalen

Die Ergebnisse der Tabelle zeigen die Wirksamkeit des Treatments in allen drei Subskalen, die sich im Falle des Posttest I in einer Zunahme der erreichten absoluten Werte und im Falle des Posttest II in den negativen Nettoeffekten zeigt. Bei der Interpretation der Daten muss bedacht werden, dass eine Vergleichbarkeit der Werte nicht gegeben ist, da die Subskalen durch unterschiedlich viele Kategorien operationalisiert sind. Aus den Ergebnissen kann daher nicht geschlossen werden, dass beispielsweise der Einfluss der MINIPHÄNOMENTA auf die experimentellen Fähigkeiten größer ist als auf die Formen spezifisch-naturwissenschaftlichen Denkens.

Damit stehen nun alle erforderlichen Kriterien für die Wirksamkeitsstufenzuweisung (vgl. Tabelle 7 und 8 in dieser Arbeit) zur Verfügung. Die folgende Tabelle fasst die Ergebnisse zusammen:

	Untersuchungsstufe I		Untersuchungsstufe II
	Posttest I	Posttest II	Posttest
Signifikanzniveau im Relation zum Prätest	hoch signifikant $p \leq 0,01$	hoch signifikant $p \leq 0,01$	hoch signifikant $p \leq 0,01$
Effektstärke	Überlappungsbereich: 42%	$\eta^2_{\text{part.}} = 0,32$	Überlappungsbereich: 47%
Umfang der Wirkung	Wirksamkeit auf drei Subskalen	Wirksamkeit auf drei Subskalen	-----
Wirksamkeitsstufe	HOCHWIRKSAM	HOCHWIRKSAM [*]	HOCHWIRKSAM

Tabelle 12: Wirksamkeit der MINIPHÄNOMENTA

Entsprechend der operationalen Kriterien ist die MINIPHÄNOMENTA zu allen Untersuchungszeitpunkten HOCHWIRKSAM auf die methodisch-formalen Fähigkeiten der Schülerinnen und Schüler.

Diskussion der Ergebnisse

Die Ergebnisse zeigen den Einfluss des Treatments auf die methodisch-formalen Fähigkeiten der Schülerinnen und Schüler zu allen Untersuchungszeitpunkten. Ähnliche Ergebnisse konnten auch im Rahmen anderer Interventionsstudien herausgearbeitet werden (zu einem Überblick vgl. Sodian, Koerber, Thoermer 2006). Schülerinnen und Schüler der oberen Klassen der Primarstufe scheinen in der Lage zu sein, derartige Fähigkeiten zu erlernen. Dabei zeigten sich in der bisherigen Literatur insbesondere Strategietrainings als wirksam (vgl. Grygier, Günther, Kircher, Sodian, Thoermer 2003). Bislang wenig bestätigt ist, dass auch offene Settings wie die MINIPHÄNOMENTA einen Beitrag zur Entwicklung der methodisch-formalen Fähigkeiten liefern können (einen ersten Beitrag in diese Richtung lieferten Sodian, Thoermer, Kircher, Grygier, Günther 2002).

Dieser Gedanke ist jedoch durchaus plausibel, denn die Konzeption der MINIPHÄNOMENTA fordert aufgrund des Fehlens inhaltlicher Informationen zur Auseinandersetzung mit formalen Inhalten und Fragen heraus. Es gibt keinerlei Versuchsanleitungen oder Erklärungen. Den Kindern steht zur Bearbeitung ihrer Ideen und Fragen nur der experimentelle Aufbau zur Verfügung. In der Ermangelung anderer Informationen scheint eine Auseinandersetzung mit methodischen Fragen notwendig, denn nur auf der Basis eines angemessenen methodischen Vorgehens kann man auch zu einem angemessenen Ergebnis kommen. Nur so kann begründet werden, warum ein Ergebnis richtig sein müsste oder ein anderes falsch ist.

⁷ Aufgrund des explorativen Charakters der Untersuchungsstufe II wird nur der globale Datentrend in Bezug auf die beiden Fragestellungen der Untersuchung ermittelt. Eine Analyse der Ergebnisse in den Subskalen erfolgt nicht.

⁸ Eventuelle Punktunterschiede zu den arithmetischen Mitteln der Tabelle 9 ergeben sich aus Rundungsdifferenzen.

Der deutliche und lang anhaltende Effekt des Treatments MINIPHÄNOMENTA auf die methodisch-formalen Fähigkeiten der Schülerinnen und Schüler lässt sich aus der Sicht des Autors auf folgende inhaltliche Argumente zurückführen: Zunächst muss bedacht werden, dass die Stationen der MINIPHÄNOMENTA den Schülerinnen und Schülern nicht nur kurz, sondern über einen erheblichen Zeitraum zur Verfügung stehen. Nach der vierzehntägigen Ausleihe erfolgt im Regelfall ein Nachbau der Stationen im kleineren Umfang, so dass das experimentelle Angebot dauerhaft bestehen bleibt. Darüber hinaus muss betont werden, dass die MINIPHÄNOMENTA aus dem Setting Schule deutlich hervorsticht. Sozialpsychologisch gesprochen: Der Stimulus MINIPHÄNOMENTA ist hoch salient. Lernen findet hier nicht im Kontext von Unterricht statt, sondern im Zusammenhang der Pausenbeschäftigung. Das situative Interesse der Schülerinnen und Schüler steht im Mittelpunkt. Sie haben die Gelegenheit selbst aktiv zu werden, Fragen zu formulieren und diese experimentell zu bearbeiten. Dabei macht ihnen die oftmals neue Erfahrung der Methode des Experimentierens Spaß. Naturwissenschaftliches Lernen wird so zu einem interessanten und amüsanten Vorgang. Über die Ebene der Konzeption hinaus, muss argumentiert werden, dass die MINIPHÄNOMENTA auch Einfluss auf den Unterricht, vor allem Sachunterricht, der Lehrerinnen und Lehrer hat (vgl. Sauer 2005; Holst 2005). Durch das Treatment kann es zu einer Stärkung der naturwissenschaftlichen, insbesondere der physikalisch-chemischen Bildung in der Primarstufe kommen. Das Experimentieren als Erkenntnisprozess, der Umgang mit Hypothesen, offenen Fragen und die Diskussionen über mögliche Erklärungen und Deutungsansätze rücken in den Vordergrund des Unterrichts. Die MINIPHÄNOMENTA wirkt dann, vermittelt über eine mögliche Unterrichtsänderung bei den Lehrerinnen und Lehrern, auf die Schülerinnen und Schüler. Als letztes inhaltliches Argument für die Wirksamkeit der MINIPHÄNOMENTA ist die intensive Einbeziehung der Eltern in die gemeinsame Arbeit des Nachbaus der Stationen zu nennen. So werden auch diese für die Thematik sensibilisiert und naturwissenschaftliche Fragen und Themen könnten auch im privaten Umfeld der Kinder verstärkt angeregt werden. Die genannten Aspekte sind spekulativ, das heißt eine empirische Überprüfung über den Umfang und die Existenz der Effekte steht noch aus. Sie plausibilisieren aber die Wirksamkeit des Treatments. Deutlich wird in diesem Zusammenhang, dass die MINIPHÄNOMENTA nicht nur direkt auf die Schülerinnen und Schüler wirkt, sondern auch vermittelt: Zum einen über eine potentielle Verbesserung des naturwissenschaftlichen Sachunterrichts, zum anderen über eine mögliche Intensivierung der Kommunikation über naturwissenschaftliche Fragen im privaten Umfeld der Schülerinnen und Schüler.

Die Diskussion der Ergebnisse der Untersuchung abschließend, soll auf die methodischen Einschränkungen der Untersuchungsstufe II hingewiesen werden. Zum einen wurde hier nur eine sehr kleine Stichprobe untersucht. Zum anderen führt der Ein-Gruppen-Prätest-Posttest-Plan zu deutlich eingegrenzten Interpretationsmöglichkeiten. Die in der Untersuchungsstufe II erzielten Ergebnisse sind daher explorativer Natur und sollten nur mit großer Vorsicht interpretiert werden. Die zentrale Schwierigkeit an diesem methodischen Vorgehen liegt in der nur begrenzten Vergleichbarkeit der beiden Gruppen der Untersuchungsstufe II zum Zeitpunkt t_0 .

Diesen Aufsatz abschließend sollen zwei Impulse aus der Untersuchung für die naturwissenschaftliche Grundbildung im Kontext des Sachunterrichts abgeleitet werden:

Hinwendung zu einem forschend-entdeckenden Unterricht: Lehrerinnen und Lehrer sind in Bezug auf naturwissenschaftliche Problemstellungen im Kontext des Sachunterrichts oft unsicher. Sachunterrichtslehrerinnen und -lehrer empfinden vielfach eine gewisse Scheu vor physikalischen und chemischen Inhalten (vgl. beispielsweise Landwehr 2002). Dies führt dazu, dass chemische und physikalische Fragestellungen im Kontext des Sachunterrichts eine nur untergeordnete Rolle spielen (vgl. Lück 2003; Sauer 2005). Die MINIPHÄNOMENTA kann nun zunächst dazu beitragen, dass eine Auseinandersetzung mit den Fragen interessant und spannend für alle Seiten sein kann. Darüber hinaus wird durch die Versuche deutlich, dass es keiner komplexen experimentellen Aufbauten bedarf. Vielmehr lässt sich ein Großteil der Versuchsaufbauten durch Materialien aus dem Supermarkt umsetzen. Über diese grundlegenden Aspekte hinaus, liefert die MINIPHÄNOMENTA aber auch konkrete methodische Hinweise. Ein zentrales Anliegen der Entwickler der MINIPHÄNOMENTA ist es, die oftmals starke Verplanung von Bildungsprozessen zu vermeiden und in der pädagogischen Arbeit zentral von den Fragen der Kinder auszugehen (vgl. Hagstedt 1999). Die Konzeption der MINIPHÄNOMENTA steht daher, im Sinne des Wagenscheinschen Ansatzes von genetischem und orientierendem Lernen, für eine Ergänzung formalisierter Ansätze (vgl. beispielsweise Lück 2003) durch offene Verfahren. Im Fokus des Unterrichts stehen damit entdeckende Lernprozesse (vgl. Hellberg-Rode 2004). Lehrerinnen und Lehrer verzichten in diesem Zusammenhang auf vorschnelle Erklärungen. Sie konzentrieren sich vielmehr auf einen diskursiven Unterrichtsverlauf, der konzeptionell in die Nähe des Ansatzes des Nachdenkens mit Kindern (vgl. Michalik, Schreier 2006; Schreier 1999; 1997) rückt. Die Ergebnisse zu den formalen Fähigkeiten der Schülerinnen und Schüler zeigen nun deutlich, dass ein solches Vorgehen fruchtbar sein kann und keinesfalls, wie – nicht zuletzt von Eltern – oft angenommen, als „experimentelle Spielerei“ abgetan werden kann.

Phänomenorientierung: Auch wenn eine alleinige Phänomenorientierung im Sinne einer oberflächlichen Auseinandersetzung mit eben diesen ohne weitere Analyse von Grygier, Günther, Kircher, Sodian, Thörmer (2003, S. 63) zu Recht kritisiert wird, unterstreichen die Ergebnisse die pädagogische Ergiebigkeit eines Aus-

gangs der naturwissenschaftlichen Grundbildung von Phänomenen. Richtig aufbereitet, bilden sie die Basis für einen tieferen Lernprozess. Dabei steht die Frage im Vordergrund, warum die Dinge so sind wie sie sind. Hierüber diskutieren Kinder gern und erfolgreich. Sie bilden eigene Erklärungen deren fachliches Niveau durchaus angemessen ist (vgl. Holst 2005). Im Kontext einer Auseinandersetzung mit den Phänomenen können die Schülerinnen und Schüler, dies zeigt diese Untersuchung, durchaus auch angemessene formal-methodische Kompetenzen entwickeln. Unterricht kann, so könnte man es zusammenfassen von Phänomenen ausgehen, sollte aber nicht bei deren Bewunderung und Bestaunung stehen bleiben, sondern eben diese Verwunderung der Kinder zum Anlass von Nachdenkgesprächen nehmen. Die experimentelle Tätigkeit wird hierbei in ein Gespräch, besser in eine Diskussion, eingebunden in deren Kontext die Lehrerin oder der Lehrer Moderatorin / Moderator ist. Sie oder er fördert eine Fragekultur, ermutigt Erklärungen zu finden und zu äußern, moderiert die Diskussion und gibt Anstöße zum vertiefenden Nachdenken.

Literaturverzeichnis

- Anderson-Inmann, L.; Horny, M. (1996): computer based concept mapping –entraicing literacy with tools for visual thinking. In: Journal of Adolescent and Adult Literacy, Nr.4
- Ansari, S. (2009): Schule des Staunens - Lernen und Forschen mit Kindern. Heidelberg: Spektrum
- Asmussen, S. (2008): Die MINIPHÄNOMENTA: Ein explorativer Überblick über die konzeptionellen, lerntheoretisch-didaktischen und bildungstheoretischen Grundlagen eines naturwissenschaftlichen Projekts für die Primarstufe. In: www.widerstreit-sachunterricht. Nr. 12
- Bortz, J. & Döring, N. (2006): Forschungsmethoden und Evaluation für Sozialwissenschaftler. Berlin: Springer
- Chalmers, A. F. (2001): Wege der Wissenschaft – Einführung in die Wissenschaftstheorie. Berlin: Springer
- Gisbert, K. (2003): Wie Kinder das Lernen lernen: Vermittlung lernmethodischer Kompetenzen. In: Fthenakis, W. E.: Elementarpädagogik nach PISA – Wie aus Kindertagesstätten Bildungseinrichtungen werden können. München: Herder
- Grygier, P.; Günther, J.; Kircher, E.; Sodian, B. & Thoermer, C. (2003): Unterstützt das Lernen über Naturwissenschaften das Lernen von naturwissenschaftlichen Inhalten im Sachunterricht? In: Cech, D.; Schwier, H-J: Lernwege und Aneignungsformen im Sachunterricht Jahresband 13: Klinkhardt
- Hagstedt, H. (1999): Läßt sich Kinderphilosophieren didaktisieren?: In: Schreier, H.: Nachdenken mit Kindern. Bad Heilbrunn: Klinkhardt
- Hellberg-Rode, G. (2004): Entdeckendes Lernen. In: Pech, D. & Kaiser, A.: Basiswissen Sachunterricht. Band 2. Baltmannsweiler: Schneider Verlag Hohengehren
- Holst, S. (2005): Entwicklung und Evaluation interaktiver Experimentierstationen – Eine Studie zur Überprüfung der Bildungswirksamkeit erfahrungsfördernder Experimentierstationen in der Primar- und Orientierungsstufe. Tönning: Der Andere Verlag
- Landwehr, B. (2002): Die Distanz von Lehrkräften und Studierenden des Sachunterrichts zur Physik. Eine empirisch-qualitative Studie zu den Ursachen. Berlin: Logos
- Lück, G. (2003): Handbuch der naturwissenschaftlichen Bildung Theorie und Praxis für die Arbeit in Kindertageseinrichtungen. Freiburg: Herder
- Michalik, K. & Schreier, H. (2006): Wie wäre es einen Frosch zu küssen? Philosophieren mit Kindern im Grundschulunterricht. Braunschweig: Westermann
- Oerter, R. & Montada, L. (1995): Entwicklungspsychologie. Weinheim: Beltz PVU
- Rost, J.; Prenzel, M.; Carstensen, C. H. & Senkbeil, M. (2004): Naturwissenschaftliche Bildung in Deutschland – Methoden und Ergebnisse von PISA 2000. Wiesbaden: VS
- Sodian, B.; Koerber, S. & Thoermer, C. (2006): Zur Entwicklung des naturwissenschaftlichen Denkens im Vor- und Grundschulalter. In: Nentwig, P. & Schanze, S.: Es ist nie zu früh! Naturwissenschaftliche Bildung in jungen Jahren. Münster: Waxmann
- Sodian, B.; Thoermer, C.; Kircher, E.; Grygier, P. & Günther, J. (2002): Vermittlung von Wissenschaftsverständnis in der Grundschule. In: Zeitschrift für Pädagogik. S. 192 - 206
- Samuelsson, I. P. & Pramling, M. A. (2007): Spielend Lernen. In: Fthenakes, W. E. & Oberhuemer, P.: Grundlagen frühkindlicher Bildung. Troisdorf: Bildungsverlag Eins
- Sauer, F. (2005): Der Einfluss offener Experimentierstationen auf das naturwissenschaftliche Lernen im Primarbereich. Tönning : Der Andere Verlag
- Schiefele, U. & Pekrun, R. (1996): Psychologische Modelle des fremdgesteuerten und selbstgesteuerten Lernens. In: Psychologie des Lernens und der Instruktion. Herausgegeben von Weinert, F. Göttingen: Hogrefe Verlag für Psychologie
- Schreier, H. (1997): Mit Kindern über Natur philosophieren. Heinsberg: Agentur Dieck
- Schreier, H. (1999): Nachdenken mit Kindern – Aus der Praxis der Kinderphilosophie in der Grundschule. Bad Heilbrunn: Klinkhardt
- Speck, J. & Acham, K. (1980): Handbuch wissenschaftstheoretischer Begriffe (Band 2: G – Q). Göttingen: Vandenhoeck & Ruprecht
- Störing, H. J. (1987): Kleine Weltgeschichte der Philosophie. Frankfurt a. M.: Kohlhammer
- Wagenschein, M. (1965): Ursprüngliches Verstehen und exaktes Denken. Stuttgart: Klett
- Wagenschein, M. (1989): Erinnerungen für Morgen - Eine pädagogische Autobiographie. Weinheim: Beltz
- Wagenschein, M. (1992): Verstehen lehren - genetisch-sokratisch-exemplarisch. Weinheim: Beltz