



Fragen und Impulse im Klassengespräch

Eine Video- und Interviewstudie in Mathematik und Deutsch an
verschiedenen Schularten

**Masterarbeit im Studiengang Forschung und Entwicklung in der
Erziehungswissenschaft**

vorgelegt von: Andrea Batzel
Nelkenstr. 23
72116 Mössingen
andreabatzel@gmx.de
Matrikelnummer: 3151202

Prüfer: Prof. Dr. Thorsten Bohl
Betreuer: Dr. Marc Kleinknecht

11. August 2010

Inhaltsverzeichnis

1	Einleitung.....	4
2	Studie: Aufgabenkultur in verschiedenen Schularten.....	6
2.1	UNTERRICHTSQUALITÄT	6
2.2	QUALITÄTSMERKMALE VON UNTERRICHT: KOGNITIVE AKTIVIERUNG UND STRUKTURIERUNG	7
2.3	AUFGABENKULTUR.....	8
3	Klassengespräch	10
3.1	FRAGEN UND IMPULSE	12
3.2	HANDLUNGSLEITENDE KOGNITIONEN.....	14
4	Deskriptive Befunde zum Unterricht an der Hauptschule und am Gymnasium.....	17
4.1	HAUPTSCHULE.....	17
4.2	GYMNASIUM	19
5	Fragestellungen	21
5.1	ZIEL DER STUDIE	21
5.2	FORSCHUNGSFRAGEN	21
6	Methode	24
6.1	DATENERHEBUNG	24
6.2	STICHPROBE.....	25
6.3	VIDEOANALYSE.....	27
6.3.1	Datenerhebung.....	28
6.3.2	Datenanalyse.....	29
6.4	INTERVIEW	36
6.4.1	Datenerhebung.....	36
6.4.2	Datenanalyse.....	38
6.5	EINSCHÄTZUNG DES KOGNITIVEN NIVEAUS DER FRAGEN	40
6.5.1	Datenaufbereitung	40
6.5.2	Datenanalyse.....	40
7	Ergebnisse.....	45
7.1	VIDEOSTUDIE	45
7.1.1	Kognitives Niveau der Fragen.....	45
7.1.2	Basiskodierung: Klassengespräch.....	48
7.1.3	Vertiefende Analyse – Handeln der Lehrkraft im Klassengespräch	49
7.2	INTERVIEWSTUDIE	50
7.2.1	Handlungsleitende Kognitionen in Bezug auf kognitive Aktivierung im Klassengespräch	50
7.2.2	Zusammenhang zwischen Vorbereitung und kognitivem Niveau der Frage	54

8	<i>Diskussion</i>	56
8.1	ZUSAMMENFASSUNG DER ZENTRALEN BEFUNDE	56
8.2	METHODISCHE ÜBERLEGUNGEN UND FOLGERUNGEN FÜR WEITERFÜHRENDE AUSWERTUNGEN.....	58
9	<i>Literatur</i>	60
10	<i>Eidesstattliche Erklärung</i>	65
11	<i>Anhang</i>	66

1 EINLEITUNG

Das Klassengespräch gilt als „typische“ Arbeitsform im deutschen Unterricht und ist seit der interkulturell vergleichend angelegten TIMS-Videostudie (1995) verstärkt in die Kritik, aber auch in den Fokus der Forschung geraten (vgl. SEIDEL ET AL. 2003a). Die starke Dominanz des Klassengesprächs zeigte sich in der Studie „Aufgabenkultur in verschiedenen Schularten“: Keine der gefilmten Unterrichtsstunden fand ohne ein Klassengespräch statt.

Die vorliegende Arbeit geht im Wesentlichen der Frage nach, welches kognitive Niveau die im Unterricht gestellten Lehrerfragen und –impulse haben. Die Studie hat somit deskriptiven Charakter und konzentriert sich auf die Beschreibung der empirisch festgestellten Unterschiede und Gemeinsamkeiten in Bezug auf das Frageverhalten der Lehrkräfte am Gymnasium, an der Realschule und Hauptschule.

Dazu wird die Arbeit zunächst in den Zusammenhang der Studie „Aufgabenkultur in verschiedenen Schularten“ (vgl. BOHL ET AL. im Druck), als deren Teilstudie sie entstanden ist, eingeordnet (Kapitel 2). Die Dimensionen kognitive Aktivierung und Strukturierung, auf denen auch die vorliegende Arbeit basiert, werden theoretisch hergeleitet und ihre Analysefunktion im Bereich der Aufgabenkulturforschung beschrieben. Kapitel 3 definiert die grundlegenden Begriffe dieser Arbeit: Klassengespräch sowie Fragen und Impulse. Darüber hinaus werden empirische Forschungsergebnisse zum Klassengespräch und zur Erforschung handlungsleitender Kognitionen zusammenfassend zitiert. Da die vorliegende Arbeit eine schulartvergleichende Perspektive einnimmt, werden in Kapitel 4 empirische Befunde zu den Polen des dreigliedrigen Schulsystems, zum Gymnasium und zur Hauptschule, dargestellt. Diese beziehen sich hauptsächlich auf die Ergebnisse der PISA-Studien. Die Aufarbeitung des aktuellen Forschungsstandes zum Klassengespräch mündet in der Formulierung der Forschungsfragen für diese Arbeit (Kapitel 5).

Im zweiten Teil der Arbeit werden zunächst die verwendeten Instrumente, ihre Entwicklung und Überprüfung beschrieben (Kapitel 6). Die vorliegende Arbeit basiert auf zwei Instrumenten (Videoanalyse und Interview), wobei die Kategorisierung und Einschätzung des kognitiven Niveaus der Fragen eine Erweiterung der Videostudie darstellt. In Kapitel 7

werden die Ergebnisse der vorliegenden Arbeit beschrieben, die in Kapitel 8 auf die Forschungsfragen rückbezogen werden. Außerdem werden abschließend das methodische Vorgehen kritisch reflektiert sowie weitere Forschungsperspektiven eröffnet.

2 STUDIE: AUFGABENKULTUR IN VERSCHIEDENEN SCHULARTEN

2.1 UNTERRICHTSQUALITÄT

Die zunehmende Autonomie der Einzelschule und der Wechsel von der input- zur output-Orientierung haben die Schulforschung verändert: Die Frage nach Schul- und Unterrichtsqualität bemisst sich nicht zuletzt daran, was bei den Schülerinnen und Schülern zu einem größeren Lernerfolg führt. Die großen Schulleistungsstudien, angefangen mit TIMSS und PISA, richteten ihren Fokus auf empirisch nachweisbare Effekte des Unterrichts, vielfältige mikrogenetische Studien haben in den Folgejahren die umfassenden Aussagen konkretisiert (z. B. die IPN Videostudie Physik oder die schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“). Videostudien bieten eine systematische Datenbasis über die Unterrichtsrealität, die unter immer neuen Aspekten ausgewertet werden kann (vgl. SEIDEL/PRENZEL 2004). Wenn im Zuge dieser Studien von Unterrichtsqualität gesprochen wird, dann werden, wie auch im Falle des Kompetenzmodells von PISA, vor allem kognitive Schülerleistungen, zunehmend auch motivationale Aspekte in den Blick genommen. Ein qualitativ hochwertiger Unterricht kann aber auch unter anderen Perspektiven bewertet werden, etwa ob er auf die Lebenslage der Schülerinnen und Schüler eingeht (vgl. BOHL 2003). EINSIEDLER (2002) versteht Unterrichtsqualität als „Bündel von Unterrichtsmerkmalen, die sich als ‚Bedingungsseite‘ (oder Prozessqualität) auf Unterrichts- und Erziehungsziele (‚Kriterienseite‘ oder Produktqualität) positiv auswirken, wobei die Kriterienseite überwiegend von normativen Festlegungen bestimmt ist und der Zusammenhang von Unterrichtsmerkmalen und Zielerreichung von empirischen Aussagen geleitet ist“ (195). Im Rahmen dieser Studie wird vor allem das Unterrichtsmerkmal kognitive Aktivierung (siehe 2.2) und seine Umsetzung im Klassengespräch untersucht. Der Begriff Unterrichtsqualität hat gemäß der Definition immer eine präskriptiv-normative Komponente. Deskriptiv verstanden erfasst Unterrichtsqualität die Wirkungszusammenhänge im Unterricht (vgl. LEUDERS 2007). Die Wirkungsweisen des Unterrichts beschreibt beispielsweise das Angebots-Nutzungs-Modell nach HELMKE (2007), das auch dieser Studie zugrundeliegt, wenn etwa angenommen wird, dass handlungsleitende Kognitionen (Lehrerpersönlichkeit) oder kognitive Aktivierung und

Strukturierung (Qualität des Unterrichts) einen Einfluss auf die Leistung der Schülerinnen und Schüler hat. Auch Merkmale der didaktischen Kommunikation, die in dieser Studie untersucht werden sollen, hängen mit der Unterrichtsqualität zusammen (vgl. LIPOWSKY ET AL. 2008).

2.2 QUALITÄTSMERKMALE VON UNTERRICHT: KOGNITIVE AKTIVIERUNG UND STRUKTURIERUNG

Für die großen Schulleistungsstudien ist also, wie in 2.1 berichtet, vielfach die Lernleistung Kriterium für guten Unterricht. Aktuelle Lernhandlungen im Unterricht werden hingegen oft nicht berücksichtigt. Die empirische Unterrichtsforschung setzt an dieser Stelle an, indem sie Unterricht unter einzelnen Qualitätsmerkmalen analysiert und diese zu Dimensionen zusammenfasst. Zwei zentrale Dimensionen sind unter anderem Strukturierung und kognitive Aktivierung (vgl. KLIEME 2006). Beide Dimensionen spielen eine entscheidende Rolle für die Unterrichtsanalyse im Projekt „Aufgabenkultur in verschiedenen Schularten“ (vgl. BOHL ET AL. im Druck).

Strukturierung wird in manchen Studien, wie auch in dieser, eher inhaltlich, in anderen Studien eher klassenorganisatorisch verstanden. Inhaltliche Strukturiertheit äußert sich beispielsweise in häufigen Zusammenfassungen oder strukturierenden Hinweisen der Lehrkraft, reduziert somit die Informationsmenge des Lerngegenstandes und erhöht so nachweisbar deren Behalten (vgl. KLIEME 2006; SEIDEL 2003a).

Kognitive Aktivierung bezieht sich auf Prozesse im Unterricht, die durch aktive Auseinandersetzung mit dem Lerngegenstand verständnisvolles Lernen unterstützen. Lernen wird in diesem Verständnis als individueller Konstruktionsprozess verstanden, der auf die aktive Auseinandersetzung des Lernenden mit den Inhalten angewiesen ist. Solche konstruktivistischen Ansätze des Lernens betonen auch stets die Vorteile des sozialen Austauschs (Ko-Konstruktion), unterstreichen somit die Bedeutung von didaktischer Kommunikation für den Lernprozess, wie sie etwa während der Klassengesprächs stattfindet (vgl. VYGOTSKIJ 1979). Denkprozesse höherer Ordnung und Verstehen sind von der kognitiven Aktivität des Lerners abhängig; sie ist für den Aufbau von Wissen und Verstehen besonders relevant (vgl. HUGENER 2008; KLIEME 2006). Kognitive Aktivierung als Dimension ist nicht an ein bestimmtes methodisches Vorgehen oder einen besonderen Instruktionsansatz gekoppelt und kann daher als Qualitätsmerkmal in ganz unterschiedlichen Unterrichtsettings

gemessen werden. Kunter et al. (2006) konnten beispielsweise in Studien zum Mathematikunterricht zeigen, dass eine stärkere kognitive Aktivierung im Unterricht durch anspruchsvollere Aufgaben zu einer höheren Lernleistung führt (vgl. JATZWAUK ET AL. 2008). Vertiefende Analysen der TIMS-Videostudie ergaben systematische Zusammenhänge zwischen dem Niveau der kognitiven Aktivierung im Unterricht und der Schülerleistung (vgl. SEIDEL ET AL. 2003a). CLAUSEN ET AL. (2003) wiederum kamen in einer Analyse von deutschem und schweizerischem Mathematikunterricht zu dem Ergebnis, dass sich kognitive Aktivierung mittels hoch-inferenter Schätzverfahren gut ermitteln ließ, worauf im Methodenteil (siehe Kapitel 6) näher eingegangen wird.

2.3 AUFGABENKULTUR¹

Die systematische Analyse der Qualität von Aufgaben ist seit den 1990er Jahren im fachdidaktischen Fokus (vgl. BLÖMEKE ET AL. 2006). Wichtig aus allgemeindidaktischer Perspektive ist die Analyse der Aufgaben im Hinblick auf vorliegende Qualitätsmerkmale, wie dies eben kognitive Aktivierung oder Strukturierung sein können. Neben dieser inhaltlichen Perspektive interessiert auch das aufgabenbezogene Handeln der Lehrkraft im Zusammenhang mit der konkreten Bearbeitung der Aufgabe im Unterricht, da Lehrpersonen durch ihre Instruktionen die Qualität einer Aufgabe verändern können. Hinweise im Vorfeld können strukturierenden Charakter haben, die Lernbegleitung zusätzlich kognitiv aktivierend sein. Stein, Grover und Hennigsen (1996) differenzieren zwischen Aufgabenstellung und Aufgabebearbeitung, BLÖMEKE ET AL. (2006) ergänzen, wie dies auch in der Studie „Aufgabenkultur in verschiedenen Schularten“ gemacht wird, die Analyse um den Auswertungs- und Reflexionsprozess. Aufgabenkultur meint also nicht nur die Analyse von Aufgaben, sondern auch deren Einbettung in den Unterrichtsprozess (vgl. JATZWAUK ET AL. 2008).

Die hier beschriebene Teilstudie „Fragen und Impulse im Klassengespräch“ nimmt als Erweiterung zur Analyseeinheit der schriftlich gestellten Aufgaben mündliche Impulse hinzu, da die engere Definition von Aufgaben (vgl. KLEINKNECHT 2010: 12) Fragen und Aufforderungen im Rahmen des Klassenunterrichts nicht berücksichtigt. Wie sich das

¹ vgl. KLEINKNECHT 2010; BOHL ET AL. 2010

Klassengespräch sowie mündlich gestellte Fragen und Impulse in die Studie einpassen, wird im folgenden Kapitel beschrieben.

3 KLASSENGESPRÄCH

Klassengespräche waren bisher nur indirekt in der Phase der Aufgabenstellung und während der Besprechung der Aufgabe im Fokus der Analyse und wurden über manche Items im hoch-inferenten Verfahren eingeschätzt. Nun sollen sie einer gezielteren Analyse unterzogen werden.

LOWYCK (1976) unterscheidet zwischen dem Klassengespräch, worunter er den partnerschaftlichen Austausch zwischen Lehrkraft und Schülerinnen und Schülern über ein eingebrachtes Thema versteht, sowie dem Unterrichtsgespräch, unter das er von der Lehrkraft als Gesprächsleitung über Fragen gesteuerte, zielgerichtete Gespräche fasst. Diese Unterscheidung scheint heute veraltet: Die Begriffe Klassengespräch und Unterrichtsgespräch werden in Praxis und Theorie synonym verwendet. Gebräuchlicher für Lehrkräfte in den Interviews erschien der Begriff Klassengespräch, weshalb er in dieser Studie verwendet wird. Unter Klassengespräch werden in dieser Studie Phasen der didaktischen Kommunikation im Plenum zwischen Lehrkraft und Klasse verstanden, die durch Fragen und Impulse der Lehrkraft gesteuert werden; es handelt sich also um eine lehrerzentrierte Sozialform. Sie dient meist der Vor- und Nachbereitung von Schülerarbeitsphasen.

Die TIMS-Videostudie 1995 belegt in der vergleichenden Analyse das typische Skript des deutschen Mathematikunterrichts: Im fragend-entwickelnden Klassengespräch werden neue Sachverhalte erarbeitet. Typisch dabei ist, dass die Lehrkraft ihre Ausführungen mit einer relativ komplexen anspruchsvollen Aufgabenstellung beginnt und diese im Klassengespräch mittels geschlossener Fragen von den Schülerinnen und Schülern bearbeiten lässt (vgl. LEUDERS 2007; SEIDEL/PRENZEL 2004; SEIDEL ET AL. 2003a; KLIEME 2002). Dies ist kein singuläres Merkmal des deutschen Mathematikunterrichts: Für den Physikunterricht ermittelten SEIDEL & PRENZEL (2004) eine durchschnittliche Klassengesprächsdauer zwischen 13.0 und 13.5 Minuten pro Unterrichtsstunde. Je nach Thema nahm das Klassengespräch bis zu 80% der Unterrichtszeit ein und spielte konstant über alle erfassten Themen eine bedeutende Rolle. Die Schülerinnen und Schüler waren dabei überwiegend nur Stichwortgeber oder wurden

von der Lehrkraft in ihren Beiträgen ergänzt. Somit bestätigt sich die dominierende Rolle der Lehrkraft im Klassengespräch.

Als Indikator für das kognitive Niveau des Unterrichts ermittelten LIPOWSKY ET AL. (2008) die Länge der Schülerbeiträge, die wiederum mit dem Niveau und der Häufigkeit der Lehrerfragen zusammenhängt. Angenommen wurde, dass komplexe, anspruchsvolle Lehrerfragen mit längeren Schülerantworten einhergehen, während Fragen zu deklarativem Wissen kurz beantwortet werden. Diese Vermutung bestätigte sich: Die kognitive Aktivierung und die mittlere Länge der Schülerbeiträge pro Klasse im Klassengespräch korrelieren (.47). Empirisch belegt sind auch positive Effekte anspruchsvoller Unterrichtsgespräche, die die Lernenden durch Fragen nach Beziehungen und Lösungsmethoden oder durch das Einfordern von Argumenten oder Erklärungen in kognitiv aktivierender Weise am Gespräch beteiligen (vgl. PAULI ET AL. 2008). SEIDEL (2003a) konnte umgekehrt zeigen, dass Anforderungen im Klassengespräch, die nur darauf zielen, Begriffe und Fakten zu nennen, bei den Schülerinnen und Schülern nicht zu einer vertieften Auseinandersetzung beitragen und somit nicht den Lernerfolg unterstützen. Lernförderliche Aspekte von kognitiv anspruchsvollen Lehrerfragen konnten in einer Metaanalyse von Gayle, Preiss & Allen (2006) nachgewiesen werden (vgl. PAULI/LIPOWSKY 2007).

Insgesamt jedoch haben Fragen, die auf eine kurze Antwort oder auf die Nennung von Fakten zielen älteren wie jüngeren Studien zufolge einen großen Stellenwert im Klassengespräch: In einer Videostudie zum Physikanfangsunterricht wurden 34% Kurzantworten ermittelt bei gleichzeitiger Dominanz der Sprechanteile der Lehrkraft am Unterrichtsgespräch (vgl. SEIDEL im Druck). Eine ältere Studie von Gallagher (1965) ergibt einen Anteil von mehr als 50% Reproduktionsfragen, wobei es unerheblich war, ob eine erfahrene oder unerfahrene Lehrkraft die Fragen stellte. LOWYCK (1976) kommt in der Zusammenfassung weiterer Studien zu dem Schluss, dass 60% der Fragen Reproduktionsfragen sind, 20% auf die Initiierung eines Denkprozesses bei den Schülerinnen und Schülern zielen und 20% Verfahrensfragen sind.

Fraglich ist an dieser Stelle, ob gerade kognitiv aktivierende Fragen, die für den Lernprozess wertvoll sind, wirklich alle Schülerinnen und Schüler erreichen oder nur diejenigen, die sich

am Gespräch aktiv beteiligen. Inagaki et al. (1998) zeigten anhand einer Stichprobe von elf japanischen Klassen, dass schweigende Schülerinnen und Schüler ebenso ein Verständnis für ein mathematisches Konzept entwickeln konnten wie diejenigen, die sich an der Diskussion der Lösungsvorschläge im Klassenverband beteiligten (vgl. PAULI/LIPOWSKY 2007). Allerdings müssen hierbei auch kulturelle Unterschiede und die höhere Disziplin japanischer Schülerinnen und Schüler mitbedacht werden.

Die aktive Partizipation an den Wissensbildungs- und Problemlöseprozessen ist ein wichtiger Aspekt kognitiver Aktivierung (vgl. PAULI/LIPOWSKY 2007). Daher werden Fragen und Impulse in dieser Studie als Angebot an die Lernenden verstanden und die Intention der Lehrkraft, die sie mit der jeweiligen Frage verbindet, analysiert und eingeschätzt.

3.1 FRAGEN UND IMPULSE

Die Qualität des Klassengesprächs bemisst sich, wie oben beschrieben wurde, vor allem an der Qualität der Lehrerfragen und Impulse, deren kognitives Niveau im Rahmen dieser Studie eingeschätzt werden soll. Fragen und Impulse stehen allerdings nicht wie Aufgaben in schriftlicher Form unmittelbar zur Analyse zur Verfügung, sondern müssen erst aus dem Transkript des Klassengesprächs gewonnen werden. JATZWAUK (2007) ermittelt die Analyseeinheit „Aufgabe“ über Operatoren. „Operatoren sind Handlungsanweisungen bzw. allgemeine Handlungsprogramme, die in Form von Verben in der Imperativform, z. B. ‚Nenne ...‘ oder äquivalenten Formulierungen in Form von Substantiven oder Fragewörtern, z. B. ‚Beobachtung: ...‘ oder ‚Welche ...?‘ vorkommen“ (JATZWAUK ET AL. 2008: 267). Während Aufgaben nach dieser Methode in der Studie „Aufgabenkultur in verschiedenen Schularten“ identifiziert und jeweils analysiert wurden, eignet sich dieses Verfahren zur Identifizierung von „Operatorfragen“ nicht, da Lehrkräfte ihre formulierten Fragen beispielsweise durch eine weitere, direkt angehängte Frage verändern: „Wie seid ihr bei der Lösung der Aufgabe vorgegangen, was war dein erster Schritt?“ Nach dem Prinzip der Operatorfrage enthält dieser Satz zwei Fragen (1. Wie seid ihr bei der Lösung der Aufgabe vorgegangen?, 2. Was war dein erster Schritt?), reagieren wird der aufgerufene Schüler bzw. die Schülerin nur auf den zweiten Teil der Frage und den ersten Schritt beschreiben.

In der schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“ (vgl. PAULI 2006) wird die Art der Fragen nach geschlossener

und offener Frage unterschieden. Das kognitive Niveau der Fragen wird mithilfe von Kategorien und Unterkategorien eingeschätzt, die berücksichtigen, dass es sich ausschließlich um Mathematikstunden zum Satz des Pythagoras oder zu Textaufgaben handelte. Eine derart enge Eingrenzung auf ein Fach und zwei Themen und die damit verbundene exakte Beschreibung der Kategorien ist angesichts der Vielfalt der videographierten Stunden dieser Studie nicht möglich.

Da kognitive Aktivierung als ein Qualitätsmerkmal von Unterricht zentral für diese Studie ist, erscheint die Bloomsche Lernzieltaxonomie (vgl. BLOOM 1972) hilfreich, um die Fragen nach ihrem kognitiven Niveau zu kategorisieren: Fragen nach Wissen aktivieren weniger kognitiv als beispielsweise Verständnisfragen oder Anwendungsfragen. Analyse, Synthese und Bewertung als kognitive Anforderung bewirken schließlich die größte gedankliche Auseinandersetzung mit einem Lerngegenstand (vgl. JATZWAUK 2007: 58). Dies passt sich ein in das Modell der Intelligenzstruktur von Guilford, in dem er zwischen Cognition (Erkennen und Wiedererkennen), Memory (Erinnern), Divergent Production (Divergente Denkprozesse, die auf mehrere Lösungen zielen), Convergent Production (Konvergente Denkprozesse, die durch Schlussfolgerung auf eine richtige Lösung zielen) und Evaluation (Bewertung) unterscheidet (vgl. JATZWAUK 2007: 60). Geschlossene Fragen können lediglich Denkprozesse des (Wieder-)Erkennens oder Erinnerns anstoßen; um Lernende zur Bewertung eines Sachverhaltes anzuregen, bedarf es offener Fragen. Die Einteilung von Bloom war Grundlage bei der zweiten Überarbeitung des Kodiermanuals.

In der IPN Videostudie Physik wird ebenfalls das kognitive Niveau der Lehrerfragen kodiert (vgl. KOBARG/SEIDEL 2003: 169ff.). Die Kategorien basieren auf der Einteilung von NIEGEMANN & STADLER (2001: 177ff.) in

- Fragen, die kein Lernen intendieren
- reine Reproduktionsfragen, d. h. Fragen, die eine Wiederholung von Lehrinhalten der aktuellen oder auch vorangegangenen Unterrichtsstunden intendieren
- Kurzantwortfragen, zu denen beispielsweise Verifikations- oder Disjunktionsfragen zählen
- Langantwortfragen, zu denen beispielsweise Definitions- oder Beispielfragen gehören

- Deep-reasoning-Fragen, zu denen u. a. Urteilsfragen zählen

Niegemann & Stadler orientieren sich an der Lernzieltaxonomie von Bloom: Die Reproduktionsfragen zielen auf Wissen ab, Kurzantwortfragen auf das Anwenden von Wissen, das Verstehen impliziert. Langantwortfragen können der Analyse und Synthese zugeordnet werden, Deep-reasoning-Fragen schließlich sind unter anderem mit einer Bewertung verbunden.

Bloom betont mehrfach, dass die Grenzen zwischen den Kategorien schwer zu ziehen sind, da jede Stufe eine Voraussetzung für die nächste darstellt, die diese dann auch beinhaltet. Zudem können zwei Schüler die gleiche Aufgabe lösen, der eine reproduktiv, während der andere allgemeine Prinzipien zur Lösung heranziehen muss. Was reproduktiv ist, hängt vom Erfahrungshintergrund des Schülers bzw. der Schülerin ab. Das gleiche Problem entsteht beim Einschätzen der Frage, selbst wenn die Klasse als Einheit betrachtet wird, die also einheitliche Voraussetzungen mitbringt. Die Fragen können also nicht zweifelsfrei einer Kategorie zugeordnet werden, da sich der Denkprozess des Schülers bzw. der Schülerin der direkten Analyse entzieht.

Als Kriterium zur Kategorisierung von Lehrfragen werden die Merkmale der von der Lehrkraft beabsichtigten (nicht der gegebenen) Antwort herangezogen, da sie Aufschluss über das kognitive Aktivierungspotenzial der Frage oder des Impulses geben. Wie auch bei der Aufgabenanalyse wird nicht die Wirkung, sondern die Intention analysiert und bewertet.

3.2 HANDLUNGSLEITENDE KOGNITIONEN

In der Literatur zur Überzeugung von Lehrkräften herrscht bislang kein Konsens über die verwendeten Begrifflichkeiten (vgl. DUBBERKE ET AL. 2008). Verwendet werden unter anderem die Begriffe „subjektive Theorien“ und „handlungsleitende Kognitionen“, die im Folgenden näher beschrieben und gegeneinander abgrenzt werden. Zudem werden Ergebnisse aus aktuellen Studien zitiert.

Subjektive Theorien sind in ihrer Struktur wissenschaftlichen Theorien ähnlich, genügen aber nicht deren Gütekriterien (Systematik, Explizitheit, Falsifizierbarkeit usw.). Darüber hinaus sind sie nicht auf der Grundlage wissenschaftlicher Überprüfungen entstanden, dienen aber dennoch zur Erklärung und Vorhersage und leiten insofern, oft stärker als wissenschaftliche

Theorien, das Handeln (vgl. HELMKE 2007; MÜLLER 2004). Sie „sind relativ stabile Strukturen (mentale Repräsentationen), die aber durch Erfahrung veränderbar sind. Damit unterscheiden sie sich von momentanen, bewussten Kognitionen, die lediglich aktualisierte Manifestationen Subjektiver Theorien oder anderer Wissensbestände sein können“ (MÜLLER 2004: 22). Subjektiver Theorien sind teilweise implizit, teilweise bewusst und mitteilbar und entstehen in einem fortlaufenden Sozialisationsprozess. Sie steuern vielfach gar nicht bewusst das Handeln der Lehrkräfte (vgl. HELMKE 2007; MÜLLER 2004). Groeben et al. (1988) definieren demgegenüber handlungsleitende Kognitionen als Handlungsregeln, die aus subjektiven Theorien abgeleitet werden (vgl. MÜLLER 2004).

Interessant für die Unterrichtsforschung ist nun, wie sich diese individuellen Überzeugungen und subjektiven Interpretationen auf das Handeln der Lehrkraft etwa beim Formulieren von Fragen im Klassengespräch auswirken. Im Forschungsprogramm Subjektive Theorien (Groeben et al. 1988) wurde die prognostische Validität der erhobenen Subjektiven Theorien untersucht. Dazu wurden aufgrund der subjektiven Theoriestruktur individuelle Prognosen für konkrete zukünftige Unterrichtssituationen erstellt und in der Realität überprüft. „Der Grad an Konsistenz lag zwischen 0.67 und 0.75. Inkonsistent verhalten sich Lehrer vor allem dann, wenn sie sich stark ärgern oder in der Verfolgung ihres Handlungsziels behindert fühlen“ (MÜLLER 2004: 31). Die Forschung zum Persönlichkeitsparadigma beschäftigt sich mit Studien, die die Persönlichkeits- und Charaktereigenschaften erfolgreicher Lehrkräfte erheben. Zwischen den Persönlichkeitsmerkmalen der Lehrkräfte und dem Lernerfolg der Schülerinnen und Schüler werden uneinheitliche oder schwache Korrelationen berichtet. Dennoch gibt es einzelne Merkmale wie die Lehrerselbstwirksamkeit oder handlungsbezogener Optimismus, denen ein Einfluss auf die Lehrer-Schüler-Interaktion nachgewiesen werden kann, womit sie sich indirekt auf die Schülerleistungen auswirken können (vgl. LIPOWSKY ET AL. 2005). Staub und Stern (2005) konnten belegen, dass die Variation in der Mathematikleistung von Zweit- und Drittklässlern durch die Unterschiede in den Überzeugungen der Lehrkräfte erklärt werden können, womit jedoch noch nicht untersucht ist, wie die Überzeugungen wirken (vgl. DUBBERKE ET AL. 2008). Neuere Forschungsergebnisse zum Einfluss des Handlungswissens von Lehrerinnen und Lehrern auf die Unterrichtspraxis zeigen aber auch, dass Wissen und beobachtbares Handeln oft nicht

übereinstimmen und stattdessen auf Handlungsrouinen zurückgegriffen wird. MÜLLER (2004) bestätigt diese Ergebnisse in seiner Studie mit einer geringen Übereinstimmung zwischen den Vorstellungen vom eigenen und tatsächlichen Handeln.

Abschließend stellt sich die Frage, ob handlungsleitende Kognitionen überhaupt nachträglich rekonstruiert werden können. 1977 wurde dies von Nisbett und Wilson in Zweifel gezogen. Eine Reihe neuerer Untersuchungen zeigt jedoch, dass unter anderem mit Hilfe von Interviews Lehrerkonzeptionen identifiziert werden konnten (vgl. MÜLLER 2004: 31). Bei der Rekonstruktion eignet sich dabei am besten die Konfrontation mit dem eigenen videographierten Unterricht oder mit konkreten Problemsituationen.

4 DESKRIPTIVE BEFUNDE ZUM UNTERRICHT AN DER HAUPTSCHULE UND AM GYMNASIUM

Das selektive Schulsystem im Bereich der Sekundarstufe I stellt, wie im Rahmen von PISA 2000 festgestellt wurde, eine Besonderheit im internationalen Vergleich dar und lohnt daher, als Aspekt in die Studie einbezogen zu werden. Da ein wesentlicher Teil der Studie der Vergleich zwischen den drei Schularten darstellt, sollen im Folgenden die beiden Pole, Hauptschule und Gymnasium, des dreigliedrigen Schulsystems im Hinblick auf Ergebnisse aus empirischen Studien näher beschrieben werden. Entgegen der Erwartungen aufgrund der theoretischen Konzeptionen der Schularten bestehen erhebliche Überlappung zwischen den Schularten, wie die Ergebnisse von TIMSS/III zeigen: Die Leistungen der Hauptschülerinnen und Hauptschüler überlappen sich mit denjenigen der Schülerinnen und Schüler von Realschulen und Gymnasien (vgl. BOHL 2003). Damit sind die Unterschiede zwischen den Schularten eventuell geringer, als die theoretische Begründung der Dreigliedrigkeit vermuten lässt.

4.1 HAUPTSCHULE

Zur Unterrichtsrealität an der Hauptschule existieren nur wenige empirische Studien. In den meisten Veröffentlichungen zur Hauptschule geht es um deren Probleme oder gar um das Ende der Hauptschule, wobei es sich dabei eher um bildungspolitische Abhandlungen als um empirisch fundierte Ergebnisdarstellungen handelt. Wichtige Erkenntnisse bieten die Ergebnisse der PISA-Studie 2000.

Mehrere Untersuchungen zeigen bei ähnlich begabten Schülerinnen und Schülern, die unterschiedliche Schularten besuchen, mit zunehmender Verweildauer in der jeweiligen Schulart einen Schereneffekt: Kinder und Jugendliche, die eine „höhere“ Schulform besuchen, machen die größeren Fortschritte. Dies wird unter anderem mit institutionellen Effekten erklärt, da spezifische Fördereffekte der unterschiedlichen Schulformen sich unterschiedlich auswirken, wobei dieser Erklärung folgend die Förderung am Gymnasium die erfolgreichere sein muss. Ein weiterer Grund sind Kompositionseffekte: Die unterschiedliche Leistungsentwicklung ist auch auf die unterschiedliche leistungsmäßige, soziale, kulturelle

und lernbiographische Zusammensetzung der Schülerschaft zurückzuführen (vgl. TRAUTWEIN ET AL. 2007). Die Hauptschule weist dabei die kritischsten Kompositionsmerkmale auf: Ein hoher Anteil an Wiederholern, ein niedriges Leistungs- und Fähigkeitsniveau, eine Konzentration von Schülerinnen und Schülern aus bildungsfernen Familien sowie ein steigender Anteil Jugendlicher aus Elternhäusern mit besonderen sozialen und privaten Belastungen kennzeichnen die Schülerschaft laut vertiefender Analyse von PISA 2000 an der Hauptschule.

Diese Analyse ergab jedoch auch, dass es unterschiedliche Typen von Hauptschulen gibt, die unterschiedlich viele Risiko- und Belastungsfaktoren kumulieren. Baumert et al. (2006) gehen davon aus, dass in Baden-Württemberg 4,8% dem schwierigen Milieu und 61,9% dem günstigen Milieu zugerechnet werden können (vgl. TRAUTWEIN ET AL. 2007). Damit kann nicht pauschalisierend von der Hauptschule als der Problemschule gesprochen werden. 75% der Hauptschulen des schwierigen Milieus sind dadurch charakterisiert, dass der Hauptschüleranteil im Einzugsgebiet gering und der Ausländeranteil vergleichsweise hoch ist (vgl. TRAUTWEIN ET AL. 2007: 5). Die Überrepräsentation von Kindern aus bildungsfernen Elternhäusern an der Hauptschule erscheint insofern wichtig, als die PISA-Studien gezeigt haben, dass in Deutschland eine sehr hohe Korrelation zwischen sozialer Herkunft und schulischer Kompetenz besteht.

Identifizieren sich Hauptschülerinnen und Hauptschüler mit ihrer Schule, so hat dies einen deutlich negativen Effekt auf ihre Bildungsaspirationen (vgl. KNIGGE 2009). Einerseits ist also die Konzentration von Schülerinnen und Schülern mit Misserfolgskarrieren von Nachteil, da eine ungünstige Arbeitsatmosphäre entsteht, die Anstrengungsbereitschaft und Erfolgserwartungen hemmt, andererseits ergaben mehrere Untersuchungen, dass der Hauptschulbesuch positive Konsequenzen für die Selbstwirksamkeitsüberzeugungen und Interessenentwicklung hat (vgl. TRAUTWEIN ET AL. 2007). SMAXVIL (2008) ermittelte in einer Fragebogenstudie, an der 931 Schülerinnen und Schüler in Nordrhein-Westfalen teilnahmen, dass Hauptschülerinnen und Hauptschüler das signifikant positivere schulische Selbstkonzept im Vergleich zu untersuchten Gymnasiasten haben.

Von der Hauptschule als Problem- oder Risikoschule kann also pauschalisierend nicht gesprochen werden. Erfolg oder Misserfolg von Schule sind in besonderer Weise vom Lehrerhandeln im Unterricht abhängig (vgl. TRAUTWEIN ET AL. 2007: 9). Die leistungsbezogenen Ergebnisse der verschiedenen Teilleistungstudien zeigen aber auch, dass Hauptschülerinnen und Hauptschüler durchschnittlich in allen Bereichen schlechter abschnitten als die Schülerinnen und Schüler an anderen Schularten.

4.2 GYMNASIUM

Empirische Ergebnisse zum Gymnasium existieren ebenfalls kaum. KIPER (2005), die für das Gymnasium die Zuschreibung eines hohen Anspruchsniveaus konstatiert, bezieht sich in ihrem Beitrag auf die Ergebnisse der PISA-Studie, von denen einige im Folgenden zitiert werden.

Die Ergebnisse von PISA 2000 zeigen, dass Gymnasiastinnen und Gymnasiasten im Vergleich zu Schülerinnen und Schülern der anderen Schularten die höchsten Mittelwerte bei der Leistung, aber auch bei kognitiv anspruchsvollen Strategien in Bezug auf die Kenntnis und die Verfügbarkeit von Lern- und Problemlösestrategien als Voraussetzungen für selbstreguliertes Lernen aufweisen. Das Gymnasium weist von den kognitiven Grundfähigkeiten die homogenste Schülerschaft auf, was durch das Abgeben von leistungsschwächeren Schülerinnen und Schülern an die Realschule bedingt wird. Die Schülerschaft des Gymnasiums kommt aus privilegierten Elternhäusern, wie nicht nur die PISA-Studie belegt. Der höhere sozioökonomische Hintergrund ist in Deutschland wie bereits oben erwähnt ein Indikator für schulische Kompetenz. Er bedeutet in der Regel auch eine stärkere Beteiligung der Eltern, weniger Disziplinprobleme im Unterricht, bessere Beziehungsstrukturen an den Schulen und eine Fokussierung des Klimas auf höhere Leistungen (vgl. BOHL 2003). Weiterführende Auswertungen des TIMSS-Videomaterials ergaben, dass die kognitive Aktivierung im Mathematikunterricht am Gymnasium vergleichbar größer ist als an der Haupt- und Realschule.

Interessant im Rahmen dieser Studie erscheint die Feststellung, dass die vorherrschende methodische Großform am Gymnasium das entwickelnde Unterrichtsgespräch ist, das gelegentlich in Diskussionen mündet. Es ist also von einer hohen Kommunikationsdichte bei

ausgeprägter Problemorientierung sowie insgesamt von einer kognitiv akzentuierten Schulkultur am Gymnasium auszugehen (vgl. KIPER 2005).

5 FRAGESTELLUNGEN

5.1 ZIEL DER STUDIE

Die vorliegende Studie soll das Frageverhalten von Lehrkräften an Gymnasien, Realschulen und Hauptschulen im Klassengespräch beschreiben. Die Grundlage dazu stellen 33 videographierte Unterrichtsstunden sowie im Anschluss an die Aufnahme geführte Interviews mit den Lehrkräften dar. Aus den Transkripten der Klassengespräche werden Fragen und Impulse identifiziert, die hinsichtlich ihres kognitiven Niveaus eingeschätzt werden. Die Auswertung erfolgt bezogen auf die einzelne Unterrichtsstunde, schulartübergreifend in den Fächern Mathematik und Deutsch sowie gesondert für die einzelnen Schularten. Ziel ist der Vergleich der Mittelwerte und somit eine Aussage über das durchschnittliche Frageniveau in den unterschiedlichen Teilstichproben.

5.2 FORSCHUNGSFRAGEN

1. Analyse des kognitiven Niveaus der Fragen:

- Ist das durchschnittliche kognitive Frageniveau schulartspezifisch zu unterscheiden?
Aufgrund der beschriebenen Forschungsergebnisse über die Schularten Hauptschule und Gymnasium kann von einem Unterschied in der Hinsicht ausgegangen werden, dass die Fragen im Klassengespräch am Gymnasium durchschnittlich ein höheres kognitives Niveau aufweisen. Diese Annahme wird darauf zurückgeführt, dass Schülerinnen und Schüler am Gymnasium durchschnittlich leistungstärker sind als Hauptschülerinnen und Hauptschüler und dass das entwickelnde Klassengespräch als vorherrschende Sozialform beschrieben wurde. Darüber hinaus konnte für den Mathematikunterricht eine höhere kognitive Aktivierung im Gymnasialunterricht festgestellt werden. Relativierend ist zu bedenken, dass auch lernschwache Schülerinnen und Schüler kognitiv aktivierend unterrichtet werden können; dies hängt nicht vom Schwierigkeitsgrad der Fragen und Impulse ab, sondern von deren Gestaltung (z. B. Offenheit) (vgl. KLIEME 2002).
- Besteht ein Zusammenhang zwischen dem Schülerredeanteil und dem kognitiven Frageniveau?
Es konnte gezeigt werden, dass die Länge der Schülerantworten mit dem kognitiven Niveau des Unterrichts zusammenhängt. Dies soll im Hinblick darauf, dass verschiedene

Schularten und verschiedene Fächer in der Studie berücksichtigt werden, erneut untersucht werden.

- Korreliert eine höhere Fragedichte mit einem durchschnittlich niedrigeren kognitiven Niveau der Fragen?

Ausgehend davon, dass Fragen mit geringerem kognitivem Niveau auf die Reproduktion von Wissen zielen, die sich in einer kurzen Schülerantwort erschöpft, muss ein Unterricht, in dem diese Art von Fragen vorherrscht, von einer hohen Fragefrequenz geprägt sein. Fragen hingegen, die beispielsweise divergente Denkprozesse erfordern und somit die Antwort mehrerer Schülerinnen und Schüler zu einem Impuls zulassen, werden in Bezug auf ihr kognitives Niveau höher eingeschätzt.

- Unterscheidet sich die Häufigkeit der Fragen auf unterschiedlichem kognitivem Niveau an den verschiedenen Schularten?

Gemäß den bisherigen Annahmen ist davon auszugehen, dass Fragen höheren kognitiven Niveaus am Gymnasium vorherrschend sind, während für die Hauptschule davon auszugehen ist, dass Reproduktions- und Wissensfragen dominieren.

- Gibt es Fragen eines bestimmten kognitiven Niveaus, die unzureichend oder gar nicht beantwortet werden, und korreliert dies mit der jeweiligen Schulart?

Ausschlaggebend für das durchschnittliche kognitive Niveau eines Klassengesprächs sind die Fragen und Impulse, zu denen die Schülerinnen und Schüler in der Lage sind, Antworten zu geben. Selbst wenn sich das Frageniveau an den verschiedenen Schularten nicht unterscheiden sollte, ist aufgrund der anzunehmenden höheren kognitiven Voraussetzungen der Schülerinnen und Schüler davon auszugehen, dass kognitiv anspruchsvollere Fragen am Gymnasium häufiger zur Zufriedenheit der Lehrkraft beantwortet werden.

2. Hoch-inferentes Rating der Klassengespräche:

- Besteht ein Zusammenhang zwischen dem Grad der kognitiven Aktivierung des Klassengesprächs und der jeweiligen Schulart?

Es ist davon auszugehen, dass Klassengespräche, die durch kognitiv anspruchsvolle Fragen bestimmt sind, insgesamt als kognitiv aktivierender eingeschätzt werden. Da bisher angenommen wurde, dass die kognitiv anspruchsvollsten Fragen und Impulse im

Gymnasialunterricht gestellt werden, wird hier analog angenommen, dass die kognitive Aktivierung im Klassengespräch am Gymnasium am höchsten ist.

3. Analyse der Interviews:

- Besteht ein Zusammenhang zwischen den handlungsleitenden Kognitionen der Lehrkraft in Bezug auf Merkmale kognitiver Aktivierung im Klassengespräch und dem durchschnittlichen kognitiven Niveau der Fragen?

Die Forschungsbefunde zum Einfluss handlungsleitender Kognitionen auf das Handeln der Lehrkräfte differieren. Insbesondere in emotional belastenden Situationen unterscheiden sich Einstellungen und tatsächliches Handeln der Lehrkräfte. Da davon auszugehen ist, dass das Fragestellen im Klassengespräch in der Regel keine emotionale Ausnahmesituation darstellt, wird angenommen, dass handlungsleitende Kognitionen Einfluss auf das Frageverhalten der jeweiligen Lehrkraft haben.

- Stellen Lehrkräfte, die das Klassengespräch nach eigener Aussage zu Hause vorbereiten, die kognitiv anspruchsvolleren Fragen?

Das Klassengespräch gilt als wenig vorbereitungsintensive Methode (vgl. u.a. ORTH 2009). Dennoch gibt es in der an der Schulpraxis orientierten Literatur viele Ratschläge für Lehrkräfte, wie sie mit der richtigen Fragetechnik anspruchsvolle Unterrichtsgespräch leiten können (vgl. GLÄSER 2004; GROEBEN 2009). Den Ratschlägen der Praxisliteratur folgend ist davon auszugehen, dass die Vorbereitung des Gesprächs sich positiv auf das Frageniveau auswirkt.

6 METHODE²

Die vorliegende Arbeit entstand im Rahmen der deskriptiv-explorativen Studie „Aufgabenkultur in verschiedenen Schularten. Eine vergleichende Analyse von Aufgaben und Lehrerhandeln im Hauptschul-, Realschul- und Gymnasialunterricht“ (DFG-Projekt (BO 3312/1-1)), die auf die Beobachtungsschwerpunkte kognitive Aktivierung, Strukturierung und Differenzierung in Bezug auf die Planungs- und Handlungsprozesse der Lehrkraft fokussiert. Nicht einbezogen in die Analyse und Bewertung wurde die Wirkung, also der Lernprozess der Schülerinnen und Schüler (vgl. BOHL ET AL. im Druck).

6.1 DATENERHEBUNG

Die Datenerhebung für das Projekt fand im Zeitraum von November 2009 bis Februar 2010 statt. Insgesamt wurden zwölf Gymnasialstunden und elf Realschulstunden gefilmt und mit den entsprechenden Lehrkräften im Anschluss Interviews geführt. Für die Hauptschulstunden konnten Videos und Interviews aus dem Vorläuferprojekt (vgl. KLEINKNECHT 2010) per Zufallsprinzip entnommen werden.

Im Gegensatz beispielsweise zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“ wurde von jeder Lehrkraft nur eine Unterrichtsstunde aufgenommen (querschnittliches Design). Die Lehrkräfte waren dabei frei in der Wahl des Themas, wurden jedoch gebeten, eine Einführungs- oder Übungsstunde zu halten mit dem Schwerpunkt auf kognitiven Lehr- bzw. Lernzielen.

Nach der Stunde wurden die Schülerinnen und Schüler gebeten, einen Fragebogen auszufüllen, der jedoch in die Auswertung der vorliegenden Arbeit nicht einbezogen und daher an dieser Stelle nicht näher erläutert wird. Mit der Lehrperson wurde möglichst zeitnah nach der Aufnahme ein leitfadengestütztes Interview geführt, in dessen Rahmen sie einen Lehrerkurzfragebogen (vgl. IPN Videostudie Physik) ausfüllte, der Merkmale zur Beschreibung der Stichprobe enthielt. Die Analyse der Unterrichtsstunden bzw. Interviews erfolgte jeweils in zwei Auswertungsphasen.

² Eine ausführliche Darstellung der Methode findet sich in BOHL ET AL. 2010.

6.2 STICHPROBE

Für die Studie wurden Lehrkräfte der Klasse 8 an Realschulen und Gymnasien im Umkreis der Universität Tübingen gesucht. Dazu wurden die Schulleiterinnen und Schulleiter verschiedener Schulen telefonisch gebeten, Lehrkräfte aus den Fächern Deutsch oder Mathematik um die Teilnahme an der Studie zu bitten. Bei der Auswahl der Lehrkräfte auf die Merkmale Geschlecht, Berufserfahrung usw. zu achten, um eine möglichst repräsentative Stichprobe zu erhalten, erwies sich als nicht praktikabel, da es ohnehin schwierig war, die notwendige Anzahl an Lehrkräften für die Studie zu gewinnen. Mit den Lehrkräften, die sich zur Teilnahme bereit erklärten, wurde telefonisch das weitere Vorgehen besprochen, unter anderem, welche Themen geeignet wären, um die Stunde mit einer der bereits gefilmten Hauptschulstunden hinsichtlich der Merkmale Fach und Zielsetzung der Stunde zu matchen. Dieses Matching erwies sich im Verlauf der Studie als zunehmend problematisch, nicht zuletzt deswegen, weil die Lehrpläne in Klasse 8 an den unterschiedlichen Schularten im Fach Mathematik sehr unterschiedliche Themen vorsehen. So dominieren algebraische Themen in der gymnasialen Stichprobe, die in den gefilmten Real- und Hauptschulstunde nur am Rand vorkommen. Nach dem Telefonat gingen der Lehrkraft verschiedene schriftliche Unterlagen zu, die zum einen nochmals wichtige Informationen für die Lehrkraft selbst und für die Eltern enthielten und zum anderen den Datenschutz für alle Beteiligten vollumfänglich zusicherten.

Die Differenzierung in Einführungs- und Übungsstunden, die vor allem für das Fach Mathematik üblich ist (vgl. HUGENER 2008: 100ff.), wurde auch bei der Einteilung der Deutschstunden vorgenommen. Insgesamt überwiegen in beiden Fächern die Einführungsstunden, eine gleiche Verteilung der Übungsstunden in den Schularten ist jedoch gewährleistet.

Schulart	Geschlecht		Berufserfahrung (in Jahren) ³				Fach	
	männlich	weiblich	unter 10	unter 20	unter 30	30 oder mehr	Mathematik	Deutsch
Gymnasium	2	10	9	1	2	0	5	7
Realschule	2	9	5	2	0	3	6	5
Hauptschule	6	4	1	3	2	4	5	5
Gesamt	10	23	15	6	4	7	16	17

Tabelle 6.1: Stichprobe

Somit lässt sich festhalten, dass die Lehrerinnen überwiegen, was im Hinblick auf das Geschlecht zu einer unausgewogenen Stichprobe am Gymnasium und an der Realschule führt. Eine ähnliche Unausgewogenheit ist bei der Berufserfahrung in der Gymnasialstichprobe festzustellen. Hinsichtlich des Faches ist die Teilstichprobe wie auch die gesamte Stichprobe ausgeglichen.

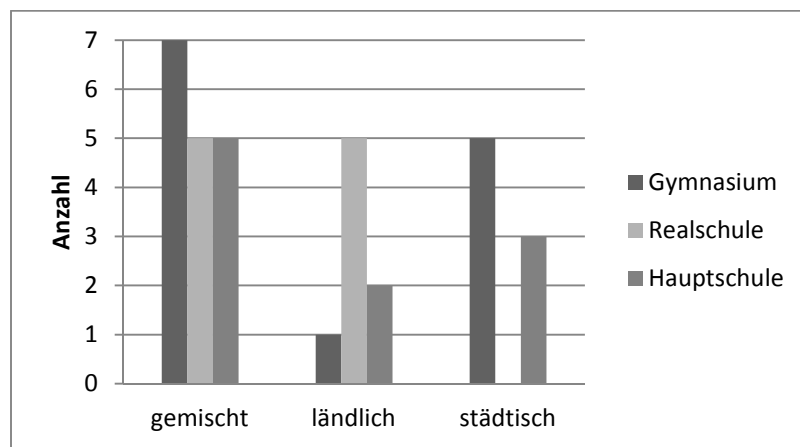


Abbildung 6.1: Einzugsgebiet der Schulen

In Bezug auf das Einzugsgebiet fällt zum einen auf, dass nur wenig ländliche Gymnasien in die Stichprobe eingehen, was mit der Verteilung der Gymnasien in eher städtischen Gebieten zusammenhängt. Zum anderen fällt auf, dass keine Realschule mit rein städtischem Einzugsgebiet in die Stichprobe eingeht.

³ Ein Wert für die Realschullehrkräfte fehlt.

Im Lehrerkurzfragebogen wurden die Lehrkräfte auch zur Repräsentativität der Stunde hinsichtlich bestimmter Merkmale befragt, die auf eine 4er-Skala⁴ beantwortet wurden. Die Ergebnisse sind in der folgenden Tabelle darstellt.

Repräsentativität	M	(SD)
Stunde	1.79	(.600)
Methodeneinsatz	1.61	(.609)
Anspruchsniveau	1.88	(.781)
Schülerverhalten	2.06	(.914)
eigene Nervosität	2.09	(.914)

Tabelle 6.2: Repräsentativität (N=33)

JATZWAUK ET AL. (2008) gehen entsprechend des Konzeptes einer relativen Konstanz von Rogosa, Floden & Willett (1984) davon aus, dass eine Unterrichtsstunde als repräsentativ für das Unterrichten der jeweiligen Lehrperson angesehen werden kann. Dies wird mit den Mittelwerten zur Repräsentativität einzelner Merkmale für die Stichprobe bestätigt.

6.3 VIDEOANALYSE

Im folgenden Abschnitt wird der Forschungsprozess für die Videoanalyse dargestellt und mit seinen Vor- und Nachteilen für die Unterrichtsforschung beschrieben. Auf Ausführungen zur Datenaufbereitung wird verzichtet und auf KLEINKNECHT (2010) und BOHL ET AL. (im Druck) verwiesen.

Videoaufnahmen wurden in der deutschsprachigen Unterrichtsforschung erstmals im Rahmen des DFG-Schwerpunktprogramms „Lehr-Lern-Forschung“ (1976-1981) verwendet (vgl. KLIEME 2006). Einzug im großen Stile erhielten Videoaufnahmen jedoch erst seit der ersten TIMS-Videostudie 1995.

Der Einsatz der Videokamera erfordert keine überdurchschnittlich technischen Fähigkeiten, macht aber dennoch ein Videotraining für die Filmenden notwendig, um wissenschaftlich verwertbare Daten zu erhalten. Die Kamera fokussiert nur einen bestimmten Ausschnitt, der vorab festgehalten werden muss, damit die Daten vergleichbar sind. Damit ist die

⁴ Skala: ‚sehr typisch‘ (1), ‚größtenteils typisch‘ (2), ‚nicht typisch‘ (3) oder ‚absolut untypisch‘ (4) bzw. ‚sehr ähnlich‘ (1), ‚ähnlich‘ (2), ‚etwas anders‘ (3) oder ‚sehr anders‘ (4) bzw. ‚gar nicht nervös‘ (1), ‚nicht sehr nervös‘ (2), ‚etwas nervös‘ (3) oder ‚sehr nervös‘ (4)

Perspektive auf Unterricht eingeschränkt (vgl. HELMKE 2007). Im Gegensatz zur teilnehmenden Beobachtung kann der Unterricht über flüchtige Beobachtungen hinaus dokumentiert sowie exakter gedeutet werden. Videodaten können mehrfach unter verschiedenen Aspekten, wie dies auch im Rahmen dieser Studie geschehen ist, ausgewertet werden. Forschungsinteressen, die erst durch eine erste Analyse des Materials entstehen, kann nachgegangen werden, ohne eine Neuerhebung notwendig zu machen.

6.3.1 DATENERHEBUNG

Für alle, die am Filmen der Unterrichtsstunden beteiligt waren (so auch die Autorin), erfolgte eine Einweisung in die Handhabung der Kamera, vor allem aber in das Kameraskript, das sich an die Richtlinien der TIMS-Videostudie 1995 anlehnt. Um Videoaufnahmen wissenschaftlich nutzen zu können, müssen bestimmte Kriterien, die der Vergleichbarkeit dienen, eingehalten werden. Im Unterschied etwa zur TIMS-Videostudie, zur IPN Videostudie Physik oder zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“ wurde nur mit einer Kamera gefilmt, die auf das Lehrerhandeln fokussieren sollte. Die Kamera blieb während des ganzen Unterrichts auf dem Stativ. Sie blieb auch dann auf die Lehrkraft gerichtet, wenn diese nicht in Interaktion mit Schülerinnen und Schülern war und etwa am Pult saß, um das aufgabenbezogene Handeln der Lehrkraft möglichst lückenlos einfangen zu können. SEIDEL ET AL. (2003b) beschreiben unter anderem zwei Problemsituationen: Während des Klassenunterrichts können mit der Lehrerkamera nicht alle Schülerinnen und Schüler, die sich am Klassengespräch beteiligen, aufgenommen werden. Dauert der Schülerbeitrag länger, so schwenkt die Kamera von der Lehrkraft zum Schüler bzw. zur Schülerin. Dieses Problem erweist sich im Hinblick auf den Schwerpunkt dieser Arbeit als sehr einschränkend, da die Interaktion der Lehrkraft im Klassengespräch oft durch die nur schwer verständlichen Schülerbeiträge beeinflusst wird. Fragen zum Aufrufverhalten der Lehrkraft oder zur Beteiligung der Schülerinnen und Schüler am Klassengespräch können nicht beantwortet werden. Verschärft wird die Problematik dadurch, dass die Gesprächsanteile der Lehrkraft durch ein an der Oberbekleidung angestecktes Funkmikrophon gut zu verstehen sind, die Beiträge der Schülerinnen und Schüler je nach Position und Raumakustik teilweise gut zu hören, teilweise unverständlich sind. Ein zweites Problem entsteht, wenn die sprechende Person und das genutzte Medium, auf das sie sich bezieht, nicht gleichzeitig gefilmt werden können. Hier sollte die Kamera sich

an der Person orientieren, jedoch das Objekt, auf das sie sich bezieht, so lange in den Blick nehmen, dass alle Informationen für die Analyse verfügbar sind.

6.3.2 DATENANALYSE

Im folgenden Abschnitt werden nur die Teile des Kodiermanuals näher beschrieben, die für die Analyse dieser Arbeit notwendig waren (vgl. BOHL ET AL. im Druck).

Entwicklung von Kodierverfahren

Zur Kodierung des Beobachtungsmaterials bedarf es eines Manuals, das theoriegeleitet in einem zyklischen Prozess zwischen Induktion und Deduktion entwickelt wird. Das Verfahren orientiert sich an der quantitativen Inhaltsanalyse (vgl. BOS/TARNAI 1999), mit dem Ziel, das Beobachtungsmaterial zu typisieren, um es der quantitativen Auswertung zugänglich zu machen (vgl. REYER 2004). Dieser zyklische Prozess besteht zunächst aus der theoretischen Eingrenzung und der Entwicklung von Dimensionen auf der Basis der Forschungsfrage. Diese Dimensionen (im hoch-inferenten Rating) sind in dieser Studie kognitive Aktivierung, Strukturierung und Differenzierung, wobei für die vorliegende Arbeit hauptsächlich die Dimension kognitive Aktivierung von Bedeutung ist. Diese Dimensionen werden in einem zweiten Schritt operationalisiert, in dem die Eigenschaften der Dimensionen bestimmt und daraus Kategorien gebildet werden, die wiederum mit Ankerbeispielen und Indikatoren konkretisiert werden. In der Überarbeitungsphase, die bereits in der Vorgängerstudie stattfand (vgl. KLEINKNECHT 2010), können weitere Beispiele als Indikatoren hinzugefügt werden. Selbst wenn bereits gültige Daten produziert werden, können sich immer wieder neue Problemfälle zeigen, die eine Überarbeitung einzelner Kategorien nötig machen. Wenn das Kategoriensystem entwickelt ist, werden die Kodierungen von jeweils zwei Kodierern bzw. Ratern durchgeführt. Die Urteile werden verglichen, abweichende Werte in der Forschungsgruppe besprochen und einem Konsensgespräch eine Einigung erzielt (vgl. KLEINKNECHT 2010: 95ff.).

Auch für Videoanalysen gelten die Gütekriterien der klassischen Testtheorie: Objektivität, Reliabilität und Validität. Objektivität meint die Genauigkeit einer Messung in dem Sinne, dass sie unabhängig von den Testanwendern (hier Kodierer bzw. Rater) ist. In diesem Fall geht die Objektivität als Exaktheit und Reproduzierbarkeit in die Reliabilität ein (vgl. REYER 2004). Reliabilität meint die Genauigkeit der Messung im Sinne der Zuverlässigkeit einer

Messmethode. Sie wird bei Videoanalysen als Interkoder-Reliabilität (niedrig-inferent) oder als Interrater-Reliabilität (hoch-inferent) angegeben. Reliabilitätsprüfungen geben an, welche Kategorien des Kodiermanuals nochmals überarbeitet werden müssen. Validität im Sinne einer Konstruktvalidität ist nicht messbar. Sie kann beispielsweise durch Methodentriangulation oder, wie im Falle des vorliegenden Manuals geschehen, in Zusammenarbeit mit anderen Forscherinnen und Forschern erreicht werden (vgl. KLEINKNECHT 2010: 96).

Aufbau des Kodiersystems

Für den Aufbau des Kodiersystems muss zwischen Sicht- und Tiefenstrukturen unterschieden werden. „Sichtstrukturen werden dabei als solche Muster der Unterrichtsorganisation verstanden, die im Verhalten der beteiligten Personen sichtbar und beobachtbar sind“ (SEIDEL/PRENZEL 2004: 179). Um die Dauer und Häufigkeit bestimmter Ereignisse zu erfassen, ist ein niedrig-inferentes Kodierungsverfahren geeignet. Es ist an beobachtbaren Verhaltensweisen ausgerichtet, die Anleitung ist definiert und klar abgegrenzt. Nachteilig an diesem Verfahren ist, dass die Validität nicht immer gewährleistet ist: Globalere Unterrichtsereignisse werden in ihrer Bedeutung vernachlässigt, da die Unterrichtsereignisse detailliert aufgegliedert werden (vgl. SEIDEL ET AL. 2003b). Die niedrig-inferente Kodierung eignet sich folglich nicht zur Beantwortung aller Fragen.

Durch die niedrig-inferente Kodierung kann die Klassengesprächsgestaltung präzise beschrieben werden, indem Sequenzen beobachtbarer Lehr- und Lernaktivität identifiziert werden, wodurch die gesamte Unterrichtsstunde in überschaubare Einheiten eingeteilt und somit das Datenmaterial für qualitative und quantitative Auswertungsschritte zugänglich gemacht wird (vgl. HUGENER ET AL. 2006: 44). Solch ein quantitativer Auswertungsschritt ist beispielsweise die Redezeit der Schülerinnen und Schüler, der der Beschreibung dient, jedoch noch keine qualitative Aussage möglich macht (vgl. NIEGEMANN & STADLER 2001; KUNTZE & REISS 2004).

Relevante niedrig-inferente Basiskodierungen für die vorliegende Arbeit sind in der folgenden Tabelle dargestellt.

Dimensionen	Kategorien
Sozialformen	Keine Klassengespräch Lehrervortrag Einzelarbeit Partnerarbeit Gruppenarbeit Mehrere Arbeitsformen gleichzeitig Andere
Art der Lehreräußerung im Unterrichtsgespräch	Frage/Aufgabe Schüleräußerung Sonstige Lehreräußerung Bedenkzeit für Schüler/innen Andere Pause Sprecherwechsel

Tabelle 6.3: Verwendete Basiskodierungen

Die Kategorien zur Dimension Sozialform wurden bereits in der Vorgängerstudie (vgl. Kleinknecht 2010) zur niedrig-inferenten Kodierung verwendet. Die Grundlage für die Beschreibung der Kategorien bilden die Ausführungen von HUGENER (2006: 55-61) und SEIDEL (2003c: 117-121). Beide lehnen sich an das Codierhandbuch der TIMSS 1999 Videostudie an. Es handelt sich um ein disjunktes Kategoriensystem, das heißt die Kodierer müssen sich für eine Kategorie je Zeiteinheit (Dimension Sozialform: 10 Sekunden) entscheiden. Grundlegend für die Kodierung der Dimension „Art der Lehreräußerung im Unterrichtsgespräch“ ist die Kategorie „Klassengespräch“. Unterrichtssequenzen, die der Kategorie Klassengespräch zugeordnet wurden, wurden mit den Kategorien der Dimension „Art der Lehreräußerung im Unterrichtsgespräch“ kodiert. MÜLLER ET AL. kodieren die Lehrer-Schüler-Interaktion im Fünf-Sekunden-Takt, um die Komplexität erfassen zu können (2006: 130). Diese Stichprobengröße stellte sich bei Probekodierungen als nicht fein genug heraus, so dass die Lehreräußerungen im Ein-Sekunden-Takt kodiert wurden. Die Formulierungen

der Kategorien zur Dimension „Art der Lehreräußerungen im Unterrichtsgespräch“ sind an KOBARG & SEIDEL (2003: 163ff.) sowie PAULI (2006: 129) angelehnt. In beiden Studien wird das Klassengespräch sehr viel detaillierter untersucht, so dass für die vorliegende Arbeit Kategorien zusammengefasst bzw. weggelassen wurden. Die Beschreibung der Kategorie 4 zeigt beispielhaft, in welcher Weise die einzelnen Kategorien operational definiert wurden.

Kategorie 4: Bedenkzeit für die Schüler/innen

Inhaltliche Bestimmung: Als Bedenkzeit werden die Anteile des Unterrichtsgesprächs kodiert, in denen die Schüler/innen Zeit zum Nachdenken über eine gestellt Frage/Aufgabe bekommen. Das namentliche Aufrufen eines Schülers/einer Schülerin ohne weitere Information fällt unter diese Kategorie, wenn vor dem Aufrufen mindestens 3 Sekunden Sprechpause waren. Nicht kodiert werden hingegen Pausen im Unterrichtsgespräch, die aus organisatorischen Gründen entstehen und von den Schüler/innen nicht zum Nachdenken genutzt werden sollen (z. B. Lehrperson sucht nach Unterrichtsmaterialien).

Beschreibung auf der Beobachtungsebene: Die Schüler/innen haben im Verlauf des Unterrichtsgesprächs eine Frage/Aufgabe erhalten, über die sie nun nachdenken sollen. Die Lehrperson wartet ab, bis sie eine Schülerin/einen Schüler aufruft. Die Bedenkzeit kann auch darin bestehen, dass die Schüler/innen von der Lehrperson aufgefordert werden, sich kurz (maximal 2 Minuten) mit dem Sitznachbarn auszutauschen.

Beispiel 6.1: Auszug aus dem Kodiermanual (Basiskodierungen)

Nicht beantwortet werden können mithilfe der niedrig-inferenten Kodierung, ob das Klassengespräch kognitive Aktivierung ermöglicht. Selbst ein hoher Redeanteil der Schülerinnen und Schüler muss nicht zwangsläufig auf kognitive Aktivierung hindeuten. Daher ist das hoch-inferente Rating unverzichtbar für qualitative Aussagen (vgl. HUGENER ET AL 2006; KUNTZE & REISS 2004).

Clausen, Reusser und Klieme halten das hoch-inferente Rating geeignet für die qualitative Bewertung von Unterrichtsereignissen wie beispielsweise das Klassengespräch (vgl. SEIDEL 2003b). Komplexe, zusammenhängende Merkmale können integriert und gleichzeitig bewertet werden, was die Einschätzung der Qualität ermöglicht. Dazu sind verstärkt interpretative Prozesse erforderlich, die auf der menschlichen Urteilsfähigkeit, verschiedene Aspekte und Kriterien gleichzeitig analysieren zu können, basiert (vgl. HUGENER ET AL. 2006). Dadurch verringert sich die Reliabilität der Kodierung. Ergebnisse von hoch-inferenten Unterrichtsbeurteilungen sind meist Skalenwerte für die untersuchten Dimensionen, die wiederum als Mittelwert über die eingeschätzten Items berechnet werden (vgl. CLAUSEN ET AL. 2003).

Das Rating wurde in Expertengruppen von zwei bis fünf Ratern je Unterrichtsstunde auf einer vierstufigen Skala vorgenommen. Die Analyseeinheit für das Klassengespräch waren

stets alle Gesprächssituationen einer Unterrichtsstunde im Gesamten, auch wenn diese zeitlich verteilt über die Unterrichtsstunde lagen.

Im Videomanual wurden exakte Definitionen der zu beurteilenden Items formuliert und mit den Raterinnen und Rater immer wieder in Konsensgesprächen besprochen. Die Items sind sowohl dem Kodiermanual zur Aufgabenanalyse entnommen, um Vergleichbarkeit und Validität herzustellen, als auch aufgrund theoretischer Annahmen über das Klassengespräch (deduktiv) sowie beobachtbarer Sequenzen in gefilmten Stunden (induktiv) ergänzt worden.

Item	Zuordnung zu Dimension
Die Lehrkraft legt Wert darauf, dass die Schüler/innen in ihren Beiträgen aufeinander eingehen.	kognitive Aktivierung
Schülerantworten werden von der Lehrkraft oft wörtlich wiederholt („Lehrer-Echo“).	
Offenheit des Klassengesprächs	kognitive Aktivierung
<i>Die Lehrkraft gibt kritisch-konstruktive Rückmeldungen.</i>	Strukturierung
<i>Die Lehrkraft gibt den Schüler/innen Denkanstöße.</i>	kognitive Aktivierung
<i>Die Lehrkraft regt die Schüler/innen an, Lösungsansätze kritisch zu überprüfen.</i>	kognitive Aktivierung
<i>Die Lehrkraft ermuntert zum selbständigen Ausprobieren.</i>	kognitive Aktivierung
<i>Die Lehrkraft verweist auf Probleme/Konflikte.</i>	kognitive Aktivierung
<i>Die Lehrkraft aktiviert das Vorwissen der Schüler/innen.</i>	kognitive Aktivierung
<i>Die Lehrkraft bezieht Fragen und Probleme ein.</i>	kognitive Aktivierung

Tabelle 6.4: hoch-inferente Items zum Klassengespräch (kursiv gedruckte Items sind dem Kodiermanual zur vertieften Analyse des aufgabenbezogenen Handelns entnommen)

Item 3: Offenheit des Klassengesprächs

Quelle: Eigene Formulierung

Die Kritik am fragend-entwickelnden Unterricht bezieht sich häufig auf ein zu eng geführtes Klassengespräch mit wenig Flexibilität für die Gedankengänge der Schüler/innen. Wünschenswert im Sinne der kognitiven Aktivierung der Schüler/innen ist ein durch offene Fragen gekennzeichnetes Klassengespräch, dessen Verlauf durch die Schülerbeiträge mitgestaltet wird, ohne dass das Ziel des Gesprächs aus dem Blick gerät.

Trifft zu:	Die Lehrkraft stellt häufig offene Fragen (Langantwortfrage oder Deep-reasoning-Frage siehe 2.1.2) und unterbricht die Schüler/innen meistens nicht in ihren Gedankengängen.
Trifft größtenteils zu:	Die Lehrkraft stellt überwiegend offene Fragen, unterbricht aber die Schüler/innen häufig in ihren Gedankengängen.
Trifft teilweise zu:	Die Lehrkraft stellt hin und wieder offene Fragen. Meistens lenkt sie durch geschlossene Fragen das Klassengespräch auf das von ihr gewünschte Ziel hin.

Trifft nicht zu: Die Lehrkraft stellt ausschließlich geschlossene Fragen. Ein Austausch zwischen den Schüler/innen ist nicht vorgesehen. Es geht vor allem um die Bewertung von richtigen und falschen Antworten.

Beispiel 6.2: Auszug aus dem Kodiermanual (Vertiefende Analysen)

Reliabilitätsprüfungen

Die Reliabilitätswerte liegen zu den von der Autorin entwickelten Teilen des Manuals nur exemplarisch oder in Auszügen vor. Für einen Großteil der Studie können die Reliabilitätswerte der Vorgängerstudie übernommen werden: Für das niedrig-inferente Verfahren der Basiskodierung wird der Cohens-Kappa-Koeffizient (κ) von Wirtz und Caspar empfohlen, der den Grad der Ähnlichkeit der Ratings misst. „In der statistischen Literatur gelten Werte von größer als .75 als Indikator für sehr gute, ein κ zwischen .60 und .75 als Maß für gute und ein Wert zwischen .40 und .60 als Beleg für noch akzeptable Übereinstimmung“ (KLEINKNECHT 2010: 104). Daneben wird die prozentuale Übereinstimmung als Prüfmaß angegeben, die laut SEIDEL (2003b) bei mindestens 85% liegen soll. Die Autorin führte an zwei Unterrichtsstunden die niedrig-inferente Basiskodierung selbst durch und glich diese mit den Kodierungen einer weiteren Raterin ab.

Stunde	Phasen		Sozialformen		Funktionen		Unterrichtsgespräch	
	PÜ (%)	Cohens κ	PÜ (%)	Cohens κ	PÜ (%)	Cohens κ	PÜ (%)	Cohens κ
GY 13	60.3	.51	89.9	.85	76.9	.66	31.4	.05
RS 10	95.8	.83	89.1	.61	88.7	.44	43.5	.13

Tabelle 6.4: Niedrig-inferente Kodierung; Übereinstimmung für N=2 Unterrichtsstunden

Es ergaben sich nur zweimal sehr gute Übereinstimmungen (RS 10 Phasen und GY 13 Sozialformen), zwei gute Übereinstimmungen (RS 10 Sozialformen und GY 13 Funktionen) sowie zwei akzeptable Übereinstimmungen (GY 13 Phasen und RS Funktionen). Damit zeigt sich, dass die Kodierung der Lehreräußerungen im Unterrichtsgespräch nur ungenügend gelungen ist, weitere Reliabilitätswerte zum Unterrichtsgespräch liegen nicht vor.

Das niedrig-inferente Kodierungsverfahren wurde in der Vorgängerstudie mit überwiegend sehr guten Übereinstimmungen ($\kappa > .75$ und PÜ > 85%) eingesetzt (vgl. KLEINKNECHT 2010: 105f.). Daher ist von ähnlichen Werten für die Basiskodierungen dieser Studie in den Dimensionen Phasen der Aufgabenbearbeitung und Sozialformen auszugehen.

Für das hoch-inferente Rating wurden zunächst, um die Validität der Items zu überprüfen, die prozentuale Übereinstimmung (PÜ) zwischen Items, die bereits für die Einführung der Aufgabe eingeschätzt werden mussten, und Items des Klassengesprächs berechnet.

Item	PÜ (%)
Lehrkraft ermuntert zum selbstständigen Ausprobieren.	65.1
Lehrkraft verweist auf Probleme/Konflikte.	58.5
Lehrkraft aktiviert das Vorwissen der Schüler/innen.	51.5
Lehrkraft bezieht Fragen und Probleme ein.	71.4

Tabelle 6.5: Prozentuale Übereinstimmung mehrfach gerateter Items

Die geringe Übereinstimmung erstaunt zunächst, da die Formulierung der Items für das Klassengespräch aus der Einführung/Erklärung der Aufgabe identisch übernommen wurde. Eine mögliche Erklärung für die niedrigen Werte könnte darin liegen, dass die Items zur Einführungsphase sich auf eine klar abgegrenzte Zeitspanne in der Unterrichtsstunde beziehen, die ohne Unterbrechung bewertet wird. Die Wertung für das Klassengespräch erfolgt nur einmal pro Unterrichtsstunde, bezieht aber alle unter Umständen verteilt über die Unterrichtsstunde liegenden Gesprächsphasen ein. Aktiviert die Lehrkraft also beispielsweise in der Einführungsphase das Vorwissen der Schülerinnen und Schüler, tut dies in der Besprechungsphase jedoch nicht, ergeben sich für die Items unterschiedliche Werte. Zu beachten ist dennoch, dass Verzerrungen in der Bewertung der Gesprächsphasen durch ihre zeitliche Trennung entstehen können.

Die Beurteilerübereinstimmung für die einzelnen Items zum Klassengespräch zeigt ebenfalls niedrige Übereinstimmungswerte (siehe Tabelle 6.6)

Bei insgesamt 10 Items für das Klassengespräch errechnen sich für

- 1 Item gute (κ zwischen .60 und .75) und
- 3 Items noch akzeptable (κ zwischen .40 und .60) Übereinstimmungen.

Item	PÜ (%)	Cohens κ
Die Lehrkraft legt Wert darauf, dass die Schüler/innen in ihren Beiträgen aufeinander eingehen.	78.8	.65
Schülerantworten werden von der Lehrkraft oft wörtlich wiederholt („Lehrer-Echo“).	69.7	.53
Offenheit des Klassengesprächs	72.7	.55
<i>Die Lehrkraft gibt kritisch-konstruktive Rückmeldungen.</i>	45.5	.18
Die Lehrkraft gibt den Schüler/innen Denkanstöße.	60.6	.39
<i>Die Lehrkraft regt die Schüler/innen an, Lösungsansätze kritisch zu überprüfen.</i>	30.3	.07
<i>Die Lehrkraft ermuntert zum selbständigen Ausprobieren.</i>	45.5	.15
<i>Die Lehrkraft verweist auf Probleme/Konflikte.</i>	57.6	.29
<i>Die Lehrkraft aktiviert das Vorwissen der Schüler/innen.</i>	36.4	.11
Die Lehrkraft bezieht Fragen und Probleme ein.	66.7	.50

Tabelle 6.6: Beurteilerübereinstimmung für einzelne Items (hoch-inferentes Verfahren); kursiv gedruckt sind Items mit sehr niedrigem Cohens Kappa-Wert

6.4 INTERVIEW

Das leitfadengestützte Interview fand in ähnlicher Form bereits in der Vorgängerstudie jeweils zeitnah nach der gefilmten Unterrichtsstunde statt. Der Leitfaden wurde gekürzt um die Frage, was die Lehrkräfte von der jeweiligen Schulart halten, und ergänzt um einen Frageteil zum Klassengespräch. Im folgenden Abschnitt wird die Darstellung auf den Frageteil zum Klassengespräch begrenzt und ansonsten auf KLEINKNECHT (2010) und BOHL ET AL. (im Druck) verwiesen.

6.4.1 DATENERHEBUNG

Das Interview gliedert sich in drei Phasen: Die erste Phase ist auf die gefilmte Stunde bezogen und fragt nach methodischen Prinzipien bei der Auswahl der Aufgaben und beim aufgabenbezogenen Handeln im Unterricht. Dazu wurde den Lehrkräften ein Schema zur Aufgabenkultur vorgelegt, das die drei Phasen im Unterricht sowie die Vorbereitung zu Hause visualisiert und dem Interviewten als Strukturierungshilfe dient. Anschließend wurden der Lehrkraft sechs Kärtchen (einfache Übungsaufgaben, anspruchsvolle Aufgaben, kleinschrittiges Erarbeiten, Anleitung und Kontrolle, Differenzierung sowie offene Lernformen) vorgelegt verbunden mit der Bitte, die Kärtchen zu kommentieren und ggf.

Beispiele aus der gefilmten Stunde zu nennen. Auch hier diene das Material wieder dazu, der Lehrkraft möglichst viel Raum zum Antworten einzuräumen und als Interviewer bzw. Interviewerin wenig Fragen stellen zu müssen. Die zweite Phase des Interviews bezog sich auf allgemeine Vor- und Einstellungen zum Lehren und Lernen an der jeweiligen Schultart, fragte also nach subjektiven Theorien und handlungsleitenden Kognitionen der Lehrkraft. In der abschließenden Phase des Interviews ging es wieder um die konkrete Unterrichtsstunde, genauer um die Gesprächsphasen des Unterrichts.

Leitfrage (Erzählaufforderung)	Check Wurde das erwähnt?	Konkrete Fragen
<u>Vorbereitung des Unterrichtsgesprächs</u> (Skizze in die Tischmitte legen) Ich möchte zum Schluss nochmals auf die Unterrichtsstunde zurückkommen und den Fokus auf die Unterrichtsgespräche/das Unterrichtsgespräch lenken. Inwieweit spielen diese Gesprächsphasen in der Vorbereitung Ihres Unterrichts eine Rolle?	vorbereitende Gedanken zum Unterrichtsgespräch	Was bereiten Sie konkret vor?
<u>Strategien bei der Leitung/Moderation</u> Worauf achten Sie während des Gesprächs?	- Strategien, die der Lehrkraft explizit bewusst sind, z. B. - Einhaltung von Gesprächsregeln - Differenzierung zwischen starken und schwachen Schülern - genügend Zeit zwischen Frage und Antwort - vorbereitende „Murmelmunden“ - Zurückhaltung der Lehrkraft <u>wichtig:</u> Dies ist nicht als Checkliste zu verstehen! Nur was von der Lehrkraft selbst genannt wird, ist relevant und kann ggf. durch Nachfragen präzisiert werden.	- Warum ist Ihnen das wichtig? - Inwiefern haben Sie das in der Unterrichtsstunde umgesetzt? - Ist den Schülerinnen und Schülern diese Strategie/Methode bewusst? (→ Metagespräch über Unterrichtsgespräche)
<u>Funktion des Unterrichtsgesprächs</u> Meine letzte Frage lautet: Wozu setzen Sie Unterrichtsgespräche ein?	- Funktion von Unterrichtsgesprächen: - Erarbeitung neuer Inhalte - Austausch von Meinungen - Planungsgespräche - Schaffung kognitiver Klarheit - Präsentation/Besprechung von Arbeitsergebnissen - Metagespräche - Wiederholung	Führen Sie Unterrichtsgespräche auch mit anderen Zielen?

Tabelle 6.7: Auszug aus dem Interviewleitfaden (Phase 3)

Die Darstellung des Leitfadens in einer Tabelle wurde modifiziert von HELFERICH (2004) übernommen.

Die Fragen des Leitfadens in Bezug auf das Klassengespräch orientierten sich an der bis dahin gelesenen Literatur zum Klassengespräch, die noch vorwiegend aus normativer Praxisliteratur bestand. Erfragt werden sollte alles, was durch Unterrichtsbeobachtungen nicht zu ermitteln ist, wie Überlegungen, die im Vorfeld des Unterrichts eine Rolle spielen, sowie handlungsleitende Kognitionen, die die Gesprächsführung beeinflussen („Worauf achten Sie während des Gesprächs?“) bzw. die Intention lenken („Wozu setzen Sie Unterrichtsgespräche ein?“).

6.4.2 DATENANALYSE

Entwicklung des Kodierverfahrens

Die Auswertung der Interviews basiert auf den theoretischen Grundlagen der qualitativen Inhaltsanalyse, die sich in drei Grundformen differenziert: Zusammenfassung, Explikation und Strukturierung (vgl. MAYRING 2008). Die Interviews wurden mithilfe der strukturierenden qualitativen Inhaltsanalyse ausgewertet, wobei auch induktiv Kategorien gebildet wurden. Zentral dabei ist, dass ein Kategoriensystem erstellt wird, welches so exakt formuliert ist, dass eine eindeutige Zuordnung von Textmaterial zu den Kategorien möglich ist. Dazu werden auf der Basis der Forschungsfrage Strukturdimensionen entwickelt. Relevant für die vorliegende Arbeit sind die Dimensionen kognitive Aktivierung, Strukturierung und Differenzierung. Zu diesen Dimensionen wurden Kategorien mit Ankerbeispielen und Kodierregeln formuliert. Dieses Kategoriensystem wurde am Material erprobt, diskutiert und auf der Grundlage der Diskussion verändert. Nach einem zweiten Materialdurchlauf und einer weiteren Verfeinerung des Kategoriensystems wurden die Interviews kodiert.

Aufbau des Kategoriensystems

Variable	Code	Definition	Ankerbeispiele	Kodierregeln
Offenheit des Klassengesprächs	UgKaOffen	Das Klassengespräch wird mit offenen Fragestellungen gestaltet.	„... zum Einstieg da mit Bildern einfach eine größtmögliche Beteiligung der Schüler zu bekommen auch mit den offenen Fragestellungen am Anfang.“	
Bezug der Schülerbeiträge aufeinander	UgKaBeitrag	Die Schüler sollen in ihren Beiträgen aufeinander Bezug nehmen.	„... in denen die Schüler gegenseitig sich ihre Meinungen sagen und wo schon auch dann aufeinander Bezug genommen wird.“	

Tabelle 6.8: Auszug aus dem Interviewanalysemanual (Dimension ‚kognitive Aktivierung‘)

Reliabilitätsprüfung

Kategorie	PÜ (%)	Kategorie	PÜ (%)
<i>kognitive Aktivierung</i>		<i>Strukturierung</i>	
Offenheit des Klassengesprächs	100.0	Regeln im Klassengespräch	100.0
Bezug der Schülerbeiträge aufeinander	71.4	geschlossenes Klassengespräch	100.0
Transfergespräch	100.0	kritisch-konstruktive Rückmeldung	-
Diskussion	100.0	Struktur des Klassengesprächs	100.0
Lehrkraft als Moderator	100.0	Ergebnissicherung	100.0
Beteiligung möglichst vieler Schüler	95.8	<i>Differenzierung</i>	
Schülerfragen	100.0	Differenzierung im Klassengespräch	100.0
Ausführliche Schülerbeiträge	100.0	<i>Vorbereitung</i>	
Alltagsbezug	100.0	Verlauf des Klassengesprächs wird vermutet	100.0
		Impulse werden vorformuliert	100.0
		Keine Vorbereitung	100.0

Tabelle 6.9: Prozentuale Übereinstimmung der Kodierungen für die Variablen zum Klassengespräch

Eine intensive Kodierschulung erübrigte sich bei der Auswertung der Interviews insofern, als diejenigen, die an der Kodierung der Interviews beteiligt waren, das Manual mitentwickelten und so mit den Kategorien vertraut waren. Dies zeigt sich auch an der prozentualen

Übereinstimmung der Kodierungen, die in der folgenden Tabelle für das Klassengespräch dargestellt sind.

Es ergab sich eine durchschnittliche prozentuale Übereinstimmung von 98,1%, was ein Hinweis darauf ist, dass die Kategorien im Kategoriensystem trennscharf und exakt formuliert wurden.

6.5 EINSCHÄTZUNG DES KOGNITIVEN NIVEAUS DER FRAGEN

6.5.1 DATENAUFBEREITUNG

Für die Datenaufbereitung zur Einschätzung des kognitiven Niveaus der Fragen wurden die in der niedrig-inferenten Basiskodierung als Klassengespräch kodierten Unterrichtsphasen transkribiert. Die Lehreräußerungen wurden dabei mit den gleichen Transkriptionsregeln verschriftlicht, die auch für die Interviews galten (vgl. KLEINKNECHT 2010: 116), die Schüleräußerungen konnten nur dann transkribiert werden, wenn sie verständlich waren.

6.5.2 DATENANALYSE

Entwicklung des Kodiervfahrens

Wie in Abschnitt 3.1 bereits beschrieben, wurde zunächst in der Literatur nach bereits erprobten Kodiervfahren gesucht. In einem ersten Versuch wurden die Kategorien von NIEGEMANN & STADLER (2001) an einem Transkript erprobt. Die Autoren unterscheiden zwischen fünf Qualitätsstufen, die sie jeweils sehr genau in ihren unterschiedlichen Ausprägungen beschreiben. In Analogie zur Entwicklung des Analysemanuals für Interviews wurde versucht, Ankerbeispiele zu finden bzw. zu formulieren.

Schwierigkeiten bereitete beispielsweise die Zuordnung von Impulsen, die die Schülerinnen und Schüler zum Erklären auffordern. Für die Analyse erschien dieses System auch insofern nicht praktikabel, da es die Kategorien zu fein aufschlüsselte und die grundlegenden Gedanken hinter der Kategorisierung so aus dem Blick geraten. Zudem wurden die Begriffe Kurz- und Langantwortfragen kritisiert, da nicht die Länge der Schülerbeiträge für die Kodierung der Frage ausschlaggebend sein soll.

Fragenkategorie	Beschreibung	Beispiel
Reproduktionsfragen	Wiederholung von Lerninhalten der aktuellen oder auch vorangegangenen Unterrichtsstunde	„In der letzten Stunde ging es um den Flächeninhalt des Kreises. Wie lautet die Berechnungsformel?“
Kurzantwortfragen	Verifikationsfragen	„Ist 9 eine Quadratzahl?“
	Disjunktionsfragen	„Soll bei dieser Aufgabe der Umfang des Kreises oder Flächeninhalt berechnet werden?“
	Konzeptfragen	„Wie nennt man die Zahlen, die ein Minus als Vorzeichen haben?“
	Elementfragen	„Was sind die Eigenschaften eines Quadrates?“ (zuvor besprochen)
	Quantifizierungsfragen	„Wie viele Angaben sind notwendig, um eine Parabel zu zeichnen?“ (zuvor besprochen)
Langantwortfragen	Definitionsfragen	„Was ist also eine Quadratzahl?“ (Dies wurde zuvor im Unterricht nicht besprochen.)
	Beispielfragen	„Eine rationale Zahl ist eine Zahl, die als Bruch bzw. als periodischer oder abbrechender Dezimalbruch geschrieben werden kann. Nenne ein Beispiel.“
	Vergleichsfragen	„Welche Gemeinsamkeiten lassen sich bei beiden Gedichten feststellen?“
	Interpretationsfragen	„Formuliert eine Aussage, die aus den im Schaubild dargestellten Daten gewonnen werden kann.“
Deep-reasoning-Fragen	Antezedenzfragen	„Bei diesem Schaubild entsteht der Eindruck, als würden die Verkaufszahlen fallen. Womit hängt das zusammen?“
	Konsequenzfragen	„Wenn die Parabel um zwei Einheiten entlang der y-Achse verschoben wird, wie ändert sich dann die Funktionsgleichung?“
	Voraussetzungsfragen	„Welche Form muss die Funktionsgleichung einer Geraden haben, damit der Graph durch den Ursprung verläuft?“
	Ziel/Motiv-Fragen	„Warum versucht Hedwig, Tell zurückzuhalten?“
	Instrumentfragen	„Welche Angaben des Kreises sind notwendig, um den Flächeninhalt zu ermitteln?“
	Erwartungsfragen	„Wie ändert sich der Flächeninhalt des Kreises, wenn der Radius verdoppelt wird?“
	Urteilsfragen	„Welches Schaubild verdeutlicht am besten, dass die Verkaufszahlen gestiegen sind?“

Tabelle 6.10: Einschätzung des kognitiven Niveaus von Fragen

Daher wurden in einem zweiten Versuch die auf Niegemann & Stadler beruhenden Kategorien aus der IPN Videostudie Physik (vgl. KOBARG/SEIDEL 2003: 169-171) reduziert und aufgrund der Erfahrungen aus dem ersten Kodierdurchgang ergänzt. Zunächst stellte sich die Frage der Identifikation von Impulsen und Fragen im Transkript. In der IPN Videostudie wird dazu jeder Satz, der mit einem Fragezeichen endet, als Frage identifiziert und erst im Nachhinein auf seine inhaltliche Relevanz hin bewertet (vgl. KOBARG/SEIDEL 2003: 166f.). In der schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“ wird dieses Verfahren ergänzt, indem auch Impulse, die einen Aufforderungscharakter enthalten als Frage kodiert werden (vgl. PAULI 2006: 132). Der Autorin erschien es sinnvoller, von den Schülerbeiträgen auszugehen: Schülerbeiträge, die nicht selbst eine Frage oder einen Impuls darstellen, sind eine Reaktion auf einen anderen Redebeitrag, wenn sie inhaltlich relevant sind. Meistens reagieren Schülerinnen und Schüler mit ihren Beiträgen auf Lehrerfragen und –impulse. Ausgehend von der Antwort wird also im Transkript zurückverfolgt, worauf der Schüler bzw. die Schülerin reagiert, und, falls es sich um einen Beitrag der Lehrkraft handelt als Frage/Impuls in einer Tabelle notiert. Auch Namensaufrufe werden aufgeführt, um die Struktur von Klassengesprächen abzubilden, in denen lange an einem Impuls oder einer Frage gearbeitet wird. Nicht berücksichtigt wurde nonverbale Kommunikation, die nicht von verbaler Kommunikation begleitet war, beispielsweise das Aufrufen eines Schülers durch Kopfnicken. Darüber hinaus wurden die Beiträge der jeweiligen Lehrkraft auf Sätze untersucht, nach denen eine Pause notiert war. Diese deutet darauf hin, dass die Lehrkraft eine Frage gestellt hat und auf Antwort wartet. Im Zweifelsfall wurde das Unterrichtsvideo zur Hilfe genommen. Auch solche Fragen, nach denen kein Schüler aufgerufen wurde, also kein Schülerbeitrag erfolgte, über den die Frage rekonstruiert hätte werden können, wurden in der Tabelle notiert. Dieses Verfahren unterscheidet sich von denen in den oben zitierten Studien, da bereits die Identifikation von Fragen aus dem Transkript mit einer Interpretationsleistung verbunden ist und somit nicht als niedrig-inferentes Kodieren eingestuft werden kann. Letztlich ersetzt diese Form des Kodierens aber das Einschätzen der Fragen auf ihre inhaltliche Relevanz. Fragen wie „Wer hat das Tagebuch?“ wurden in die Bewertung der Fragen gar nicht einbezogen.

Aufbau des Kodiermanuals

Die Kategorien und ihre Beschreibungen wurden in einem Kodiermanual festgelegt.

Kategorie 3: Faktenfrage

Inhaltliche Bestimmung: Faktenfragen sind die Fragen, die auf eine kurze Antwort abzielen, aber keine Reproduktionsfragen sind. Ziel dieser Fragen ist das Nennen von Fakten oder Zahlen. Die erwartete Antwort ist dabei nur „Ja“ oder „Nein“.

Beschreibung auf der Beobachtungsebene: Faktenfragen zeichnen sich dadurch aus, dass sie nur eine kurze Antwort wie „Ja“, „Nein“ oder einen bestimmten Begriff zulassen. Ein Beispiel für eine solche Frage wäre: „Wie viel Newton erzeugen einhundert Gramm Schokolade nach euren Messungen?“

Spezifische Kodierungsregel: „Faktenfrage“ wird nur dann kodiert, wenn die erwartete Antwort keine Reproduktion von bereits erlerntem Wissen ist. Beispiele für solche „neuen“ kurzen Antworten sind zum Beispiel Ergebnisse aus Experimenten oder Aufgaben. Auch Begriffsdefinitionen oder das kurze Nennen von eigenen Vorstellungen, wie zum Beispiel auf die Frage „Woran erkennt man, dass eine Kraft wirkt?“, fallen in diese Kategorie.

Kategorie 4: Begründungsfrage (Begründungen, Erklärungen, Beispiele, Bewertungen)

Inhaltliche Bestimmung: Die Begründungsfragen sind Fragen, die eine längere Erklärung erfordern, die jedoch rezeptiv ist. Das heißt, die Antwort auf die Frage ist eine Erklärung, aber der erklärte Sachverhalt ist nicht neu, sondern den Schüler/innen bereits bekannt.

Beschreibung auf der Beobachtungsebene: Die Fragen dieser Kategorie erfordern als Antworten Beispiele, Vergleich zwischen verschiedenen Sachverhalten oder Interpretation oder Bewertungen der in Frage stehenden Sachverhalte. Ein Beispiel für eine Begründungsfrage wäre „Wo findet ihr Linsen im Alltag?“. Diese Frage erfordert die Nennung mehrerer Beispiele.

Spezifische Kodierungsregel: „Begründungsfrage“ wird immer dann kodiert, wenn die Frage auf die Erklärung oder Erläuterung eines Sachverhaltes abzielt, der in dieser oder ähnlicher Form bereits bekannt war, oder Beispiele genannt werden sollen. Begründungsfragen haben immer rezeptiven Charakter. Präsentationen von Ergebnissen aus vorangegangenen Schülerarbeitsphasen, die mit Erklärungen beispielsweise zur Vorgehensweise einhergehen, fallen ebenfalls in diese Kategorie.

Beispiel 6.2: Auszug aus dem Manual zur Kodierung des kognitiven Niveaus der Fragen

Frage/Impuls der Lehrkraft	beantwortet	Bewertung
Also, wer liest denn mal die Aufgabe vor? Sebastian	1	2
Was ist gefragt? Und das ist ja immer unsere Variable Also: $x = \dots$ Sebastian. Irgend ne Idee?	0	3
Eine Maschine. Ist das, das gefragt?	0	6
Redest mal bitte in ganzen Sätzen mit mir?	0	6
Was ist denn gefragt? Wenn Sie beide Maschinen laufen lässt, es ist nach der Zeit gefragt von beiden Maschinen. Wie ist also x definiert?	0	3
Die Frage ist: Wie lang brauchen beide Maschinen. Also ist $x = ? \dots$ Tim	0	3
Und zwar was beid(?...) beide Maschinen?	1	3
Wie lang braucht Maschine 1? eins in Stunden?... Tim.	1	3
Wie viel Drucke schafft pro Stunde schafft sie in nem Bruchteil ausgedruckt? Von dem gesamten Auftrag schafft sie in einer Stunde wieviel dann?	1	3
wie lang braucht die Maschine 2? Tamara?	1	3
Und wieviel(?...) Wieviel Drucke pro Stunde im Bruch ausgedruckt?	1	3
Beide Maschinen zusammen, das war ja unser(?...) Unsere Definition jetzt, Larissa.	1	3

Tabelle 6.11: Auszug aus der Kodiertabelle für die Einschätzung des kognitiven Niveaus der Fragen

Die Fragen wurden außerdem darauf hin bewertet, ob sie zur Zufriedenheit der Lehrkraft beantwortet wurden oder nicht. Dies erscheint insofern wichtig, als Fragen, die von den Schülerinnen und Schülern nicht so beantwortet werden können, wie die Lehrkraft dies erwartet, ihre Intention verfehlen, die mit der Kodierung eingeschätzt wurde.

Reliabilitätsprüfung

Zur Reliabilitätsprüfung wurde exemplarisch von drei Unterrichtsstunden das kognitive Niveau der Fragen von einer weiteren Person eingeschätzt. Die Auswertung wurde auf prozentuale Übereinstimmung mit der Kodierung der Autorin überprüft. Das Ergebnis dieser Überprüfung ist in der folgenden Tabelle dargestellt.

Stunde	Identifizierung der Fragen	beantwortet		Bewertung
	PÜ (%)	PÜ (%)	Cohens κ	PÜ (%)
GY 04	77.8	81.0	.44	50.0
GY 05	63.2	87.5	.33	58.3
RS 07	78.0	93.5	.80	80.4

Tabelle 6.12: Übereinstimmung der Kodierungen in Bezug auf das kognitive Niveau der Fragen

Da die die Kodierung, ob eine Frage beantwortet wurde oder nicht, auf niedrig-inferentem Niveau erfolgt, wurde für die Reliabilität auch Cohens Kappa berechnet. Für die Stunde RS 07 ergibt sich ein sehr guter Wert, für GY 04 ein akzeptabler. Der schlechte Wert für die Stunde GY 05 liegt unter anderem daran, dass nur zwei Fragen nicht beantwortet wurden, wovon eine nicht übereinstimmend bewertet wurde. Die prozentualen Übereinstimmungen für die Identifizierung der Fragen und das Einschätzung der Beantwortung der Fragen sind sehr gut bis akzeptabel. An der niedrigen prozentualen Übereinstimmung in Bezug auf die Bewertung wird deutlich, dass eine ausführliche Schulung des zweiten Kodierers hätte stattfinden müssen und dass das Kodiermanual für weitere Analysen überarbeitet werden müsste.

7 ERGEBNISSE

7.1 VIDEOSTUDIE

7.1.1 KOGNITIVES NIVEAU DER FRAGEN

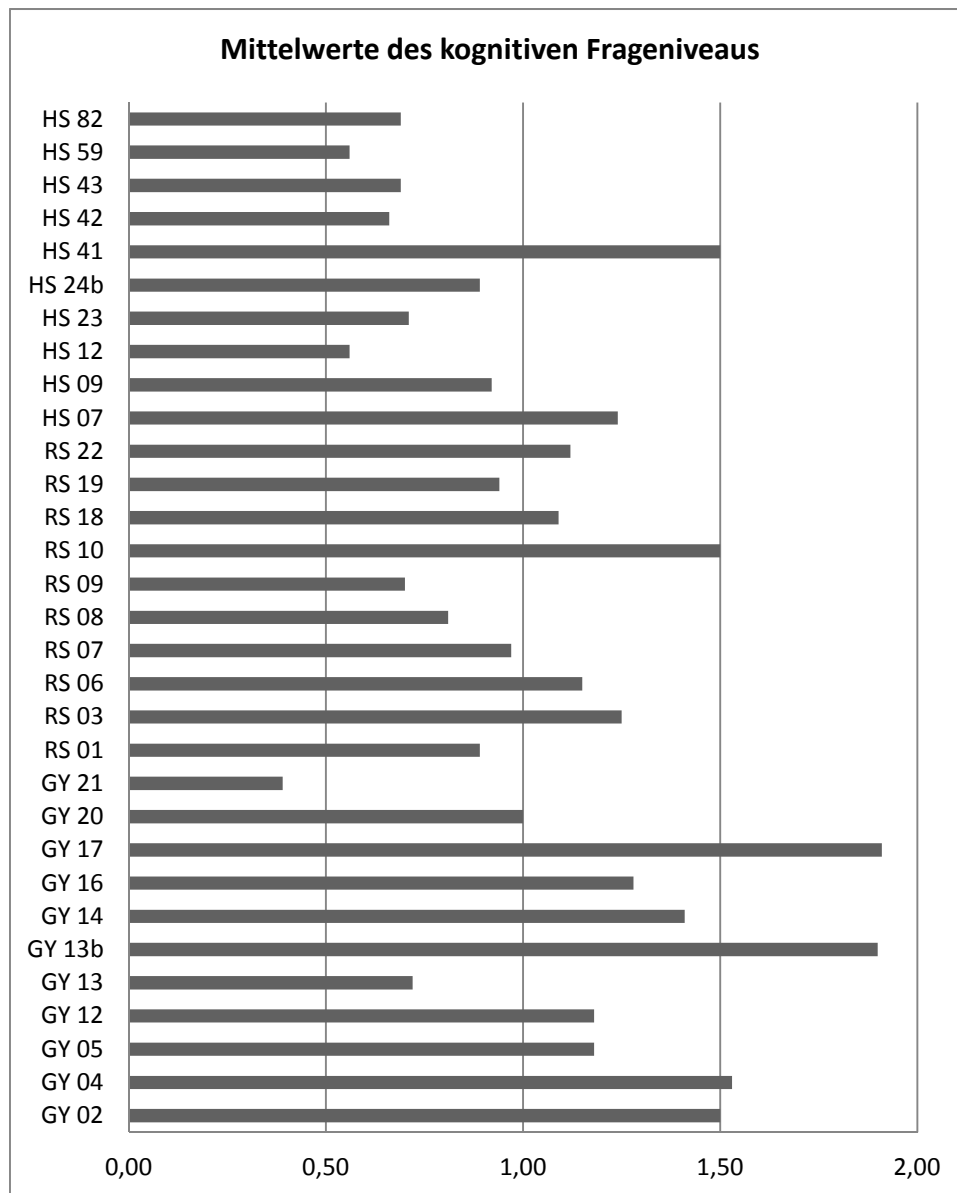


Abbildung 7.1: Mittelwerte des kognitiven Fragenniveaus pro Stunde⁵

In die Berechnung des Mittelwertes wurden nur die beantworteten Fragen einbezogen. Bei Fragen, die die Schülerinnen und Schüler nicht zur Zufriedenheit der Lehrkraft beantworten konnten, ist davon auszugehen, dass die Intention der Lehrkraft, die Schülerinnen und

⁵ Skala: ‚Reproduktionsfragen‘ (0), ‚Faktenfragen‘ (1), ‚Begründungsfragen‘ (2), ‚deep-reasoning-Fragen‘ (3)

Schüler auf einem bestimmten Niveau kognitiv zu aktivieren, nicht umgesetzt werden konnte. Die Abbildung des durchschnittlichen kognitiven Niveaus zeigt, dass es durchaus Stunden gibt, in denen das Klassengespräch aufgrund des durchschnittlichen Fragenniveaus die Schülerinnen und Schüler kognitiv aktiviert (HS 41, RS 10, GY 17, GY 13b, GY 04 und GY 02). Die meisten Unterrichtsgespräche bleiben aber hinter dem Anspruch auf kognitive Aktivierung zurück.

Eine Zusammenfassung und erneute Mittelwertberechnung nach Schularten, Fächern sowie Erarbeitungs- und Übungsstunden ergibt folgende Werte:

Teilstich- probe	Gymna- sium	Real- schule	Haupt- schule	Deutsch	Mathe- matik	Erarbei- tung	Übung
Mittel- wert (SD)	1.14 (.81)	.99 (.61)	.79 (.61)	1.16 (.76)	.87 (.66)	1.10 (.75)	.82 (.61)

Tabelle 7.1: Mittelwerte (von Teilstichproben) des kognitiven Fragenniveaus

Die Mittelwerte der Teilstichproben Gymnasium, Realschule und Hauptschule, die Mittelwerte der Teilstichproben Deutsch und Mathematik sowie die Mittelwerte der Teilstichproben Erarbeitung und Übung wurden mit T-Tests verglichen: Sie unterscheiden sich jeweils signifikant.

An den Mittelwerten wird deutlich, was in der Grafik bereits ersichtlich ist: Je höher die Schulart anzusiedeln ist, desto höher ist auch das durchschnittliche kognitive Fragenniveau. Dass sowohl beim Deutschunterricht als auch bei den Erarbeitungsstunden die Mittelwerte höher sind als bei den Mathematik- und Übungsstunden, hängt damit zusammen, dass eine geringe Korrelation (.34) zwischen den Deutsch- und Erarbeitungsstunden besteht. Das erklärt aber noch nicht, ob der Mittelwert höher ist, weil es sich um Deutschstunden handelt oder um Erarbeitungsstunden, oder ob kein Kausalzusammenhang besteht. Eine weitere Auffälligkeit besteht darin, dass die Standardabweichung steigt, wenn der Mittelwert größer wird. Das bedeutet, dass die hohen Mittelwerte am Gymnasium, im Deutschunterricht und in den Erarbeitungsstunden sich nicht aus Stunden mit einheitlich anspruchsvollen Klassengesprächen zusammensetzen, sondern dass vielmehr beispielsweise auch am Gymnasium Klassengespräche mit geringem Anspruchsniveau vorkommen (GY 21).

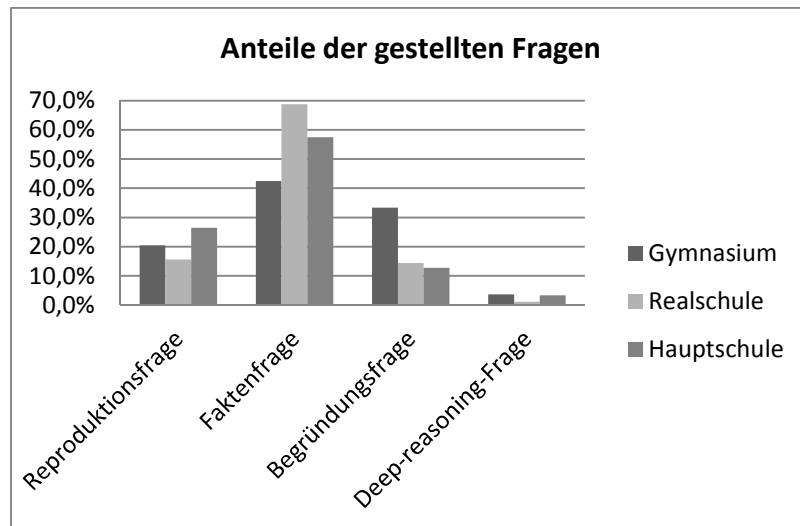


Abbildung 7.2: Anteile der gestellten Fragen

Die Verteilung der Anteile auf die einzelnen Frageniveaus unterschieden nach Schularten zeigt, dass an allen Schularten Faktenfragen überwiegen und Deep-reasoning-Fragen eine sehr geringe Rolle einnehmen. Zu beachten ist, dass hier alle Fragen analysiert wurden, die die jeweilige Lehrkraft stellte, unabhängig davon, ob die Schülerinnen und Schüler sie beantworten konnten. Begründungsfragen, die zwar reproduktiven Charakter haben, die aber dennoch zum Elaborieren beitragen, da sie beispielsweise nach Zusammenhängen oder Begründungen fragen oder die Schülerinnen und Schüler aufgefordert sind, Beispiele zu nennen, nehmen am Gymnasium im Gegensatz zu den beiden anderen Schularten eine herausragende Position ein. Der größere Anteil Faktenfragen an der Realschule geht mit einem verringerten Anteil an Reproduktionsfragen einher. Begründungsfragen werden sowohl an der Realschule als auch an der Hauptschule demgegenüber selten gestellt.

Wie bereits oben erwähnt, muss beispielsweise das häufige Stellen von Begründungsfragen am Gymnasium noch nicht dazu führen, dass dort die Schülerinnen und Schüler während des Klassengesprächs kognitiv aktiver sind. Fragen der Lehrkraft, die nicht beantwortet werden konnten, können einerseits darauf hinweisen, dass sie die Schülerinnen und Schüler überfordert haben, so dass diese gar nicht anfangen, darüber nachzudenken. Andererseits müssen kognitiv anspruchsvolle Denkprozesse nicht in eine richtige Lösung münden. Dennoch gibt der Anteil der beantworteten Fragen eher Aufschluss darüber, welche Fragen die Schülerinnen und Schüler zu zielführenden Überlegungen anregen.

Schulart	Reproduktionsfrage		Faktenfrage		Begründungsfrage		Deep-reasoning-Frage	
	nicht b. (%)	b. (%)	nicht b. (%)	b. (%)	nicht b. (%)	b. (%)	nicht b. (%)	b. (%)
Gymnasium	13.0	87.0	23.2	76.8	29.4	70.6	37.5	62.5
Realschule	21.7	78.3	31.5	68.5	32.8	67.2	20.0	80.0
Hauptschule	23.1	76.9	32.3	67.7	51.7	48.3	93.3	6.7

Tabelle 7.2: Vergleich von beantworteten und nicht beantworteten Fragen

Die Beantwortung der Frage und das Frageniveau korrelieren lediglich an der Hauptschule und dort auch nur gering (.27). Dennoch lassen sich auffällige Unterschiede und Häufigkeiten feststellen: An der Hauptschule wurden Fragen, die als kognitiv aktivierend einzustufen sind, überwiegend nicht beantwortet. Am Gymnasium sinkt der Anteil beantworteter Fragen mit zunehmendem Kognitionsniveau, bleibt aber durchgehend höher als der Anteil nicht beantworteter Fragen. Ähnliches ist für die Realschule festzustellen. Die richtige Beantwortung aller gestellten Fragen sinkt vom Gymnasium (76.3%) über die Realschule (70.0%) bis zur Hauptschule (65.6%).

7.1.2 BASISKODIERUNG: KLASSENGESPRÄCH

Die niedrig-inferente Basiskodierungen der Stunden lassen Aussagen über Häufigkeiten und Dauer bestimmter Ereignisse zu.

	Gymnasium	Realschule	Hauptschule	Deutsch	Mathematik	Erarbeitung	Übung
Anzahl der Fragen pro Minute	2.3	3.1	3.6	2.6	3.1	2.7	3.3
Mittelwert (SD)	1.14 (.81)	.99 (.61)	.79 (.61)	1.16 (.76)	.87 (.66)	1.10 (.75)	.82 (.61)

Tabelle 7.3: Fragefrequenz

Die Tabelle zeigt die durchschnittliche Fragefrequenz für die Stunden der jeweiligen Schulart bzw. des jeweiligen Faches, ergänzt um die Fragefrequenz in den Erarbeitungs- und Übungsstunden. Ein Vergleich mit dem durchschnittlichen Frageniveau zeigt, dass die Fragefrequenz steigt, wenn das kognitive Niveau der Fragen abnimmt. Zwischen der

durchschnittlichen Fragefrequenz und dem durchschnittlichen kognitiven Frageniveau besteht eine sehr hohe Korrelation (-.94**).

Darüber hinaus wurde der in der Literatur beschriebene Zusammenhang zwischen der Antwortlänge der Schülerbeiträge und dem durchschnittlichen kognitiven Niveau berechnet. Dies geschah auf der Grundlage der einzelnen Stunden, wozu zunächst aus der niedrig-inferenten Analyse der Redeanteil der Schülerinnen und Schüler im Klassengespräch ermittelt wurde. Dieser korreliert gering mit dem durchschnittlichen Frageniveau der jeweiligen Stunde (.55**). Dies bedeutet also, dass Klassengespräche, in denen die Schülerinnen und Schüler mehr Zeit für ihre Beiträge erhielten, kognitiv aktivierender waren. Dies bedeutet jedoch nicht, dass die einzelnen Schülerbeiträge länger sind. Ein hoher Schülerredeanteil kann auch entstehen, wenn viele Schülerinnen und Schüler mit kurzen Antworten ins Gespräch einbezogen werden.

7.1.3 VERTIEFENDE ANALYSE – HANDELN DER LEHRKRAFT IM KLASSENGESPRÄCH

Im folgenden Abschnitt werden Einzelergebnisse aus den Items zur kognitiven Aktivierung im Klassengespräch aus der hoch-inferenten Analyse berichtet. Die Ergebnisse müssen allerdings unter der Einschränkung der geringen Beurteilerübereinstimmung, die in Abschnitt 6.3.2 dargestellt wurde, interpretiert werden.

Item ⁶	Gymnasium	Realschule	Hauptschule
Die Lehrkraft legt Wert darauf, dass die Schüler/innen in ihren Beiträgen aufeinander eingehen.	2.2 (.94)	1.5 (.67)	1.3 (.42)
Offenheit des Klassengesprächs	2.8 (.72)	2.5 (.87)	1.8 (.13)
Die Lehrperson gibt den Schüler/innen Denkanstöße.	2.3 (.62)	1.5 (.47)	1.5 (.37)
Die Lehrkraft aktiviert das Vorwissen der Schüler/innen.	2.1 (.76)	2.3 (.75)	2.4 (.84)

Tabelle 7.4: Mittelwerte (SD) zu Einzelitems der kognitiven Aktivierung im Klassengespräch

Die Mittelwerte der Einzelitems zeigen, dass die kognitive Aktivierung im Klassengespräch am Gymnasium überwiegend höher eingeschätzt wurde als an den anderen Schularten. Der Mittelwert der Realschule liegt mit Ausnahme des Items zur Offenheit des Klassengesprächs

⁶ Skala: ‚trifft nicht zu‘ (1), ‚trifft teilweise zu‘ (2), ‚trifft größtenteils zu‘ (3), ‚trifft zu‘ (4)

dabei immer näher am Mittelwert der Hauptschule. Damit nimmt das Gymnasium in der Bewertung des Klassengesprächs in Bezug auf die kognitive Aktivierung eine Sonderstellung ein. Festgehalten werden kann außerdem, dass Klassengespräche am Gymnasium größtenteils als offen eingeschätzt werden. Alle anderen Items werden in Bezug auf die kognitive Aktivierung eher gering eingeschätzt. Auffällig ist zudem, dass sich die festgestellte Tendenz bei der Aktivierung des Vorwissens umkehrt: Hier ist der Mittelwert an der Hauptschule am höchsten. Dies könnte mit der großen Bedeutung von Reproduktions- und Faktenfragen an dieser Schulart zusammenhängen.

7.2 INTERVIEWSTUDIE

7.2.1 HANDLUNGSLEITENDE KOGNITIONEN IN BEZUG AUF KOGNITIVE AKTIVIERUNG IM KLASSENGESPRÄCH

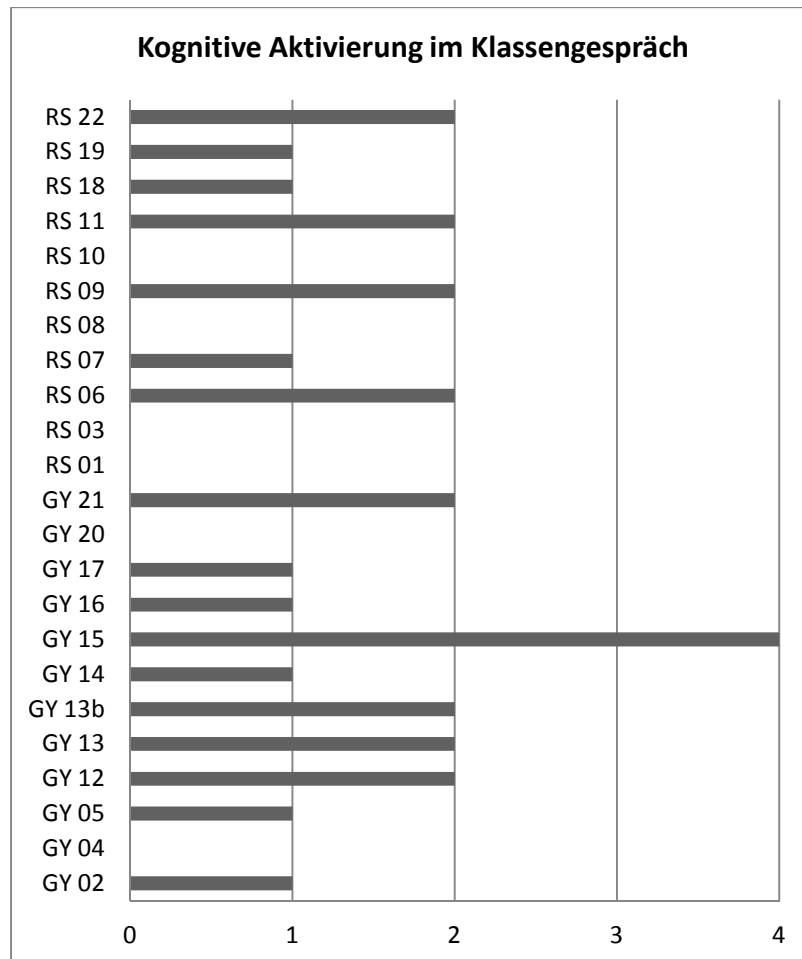


Abbildung 7.3: Kognitive Aktivierung im Klassengespräch aus Sicht der Lehrkräfte

Im Interview wurden die Lehrkräfte gefragt, worauf sie bei der Vorbereitung und während des Klassengesprächs Wert legen und wozu sie Klassengespräche einsetzen. Die Antworten der Lehrkräfte wurden nach Aussagen zur kognitiven Aktivierung, Strukturierung und Differenzierung analysiert. Für die Auswertung wurde dazu für jede Variable die häufigste Nennung über alle Interviews mit 100% gleichgesetzt. Dieser Wert ist nicht absolut, sondern relativ zu interpretieren: Über die Stichprobe gesehen, war dies die häufigste Nennung. Die Prozentwerte zu den Items der kognitiven Aktivierung wurden addiert und als Summe auf einer 4er-Skala linear verteilt.

In die Analyse konnten nur die Interviews mit den Gymnasial- und Realschullehrkräften einbezogen werden, die Hauptschulstunden und die dazugehörigen Interviews wurden aus der Vorgängerstudie (vgl. KLEINKNECHT 2010) übernommen wurden. Der Interviewleitfaden der Vorgängerstudie sah jedoch keinen Frageteil zum Klassengespräch vor.

Im Hinblick auf den Zusammenhang zwischen handlungsleitenden Kognitionen und dem Handeln in konkreten Unterrichtssituationen wurde verglichen, inwiefern Lehrkräfte, die nach eigener Aussage auf Merkmale der kognitiven Aktivierung im Unterricht achten, kognitiv anspruchsvollere Fragen stellen. Auch auf der Skala des Fragenniveaus kann kein absoluter Wert für kognitive Fragen angegeben werden. Daher werden die Mittelwerte der Gymnasial- und Realschulstunden linear analog zum Vorgehen bei der Darstellung der Interviewaussagen auf einer Skala von 0 bis 4 verteilt.

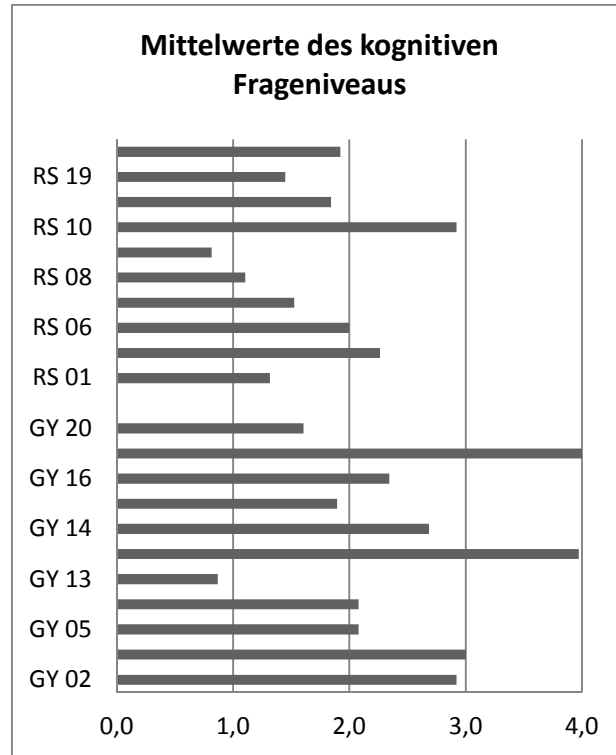


Abbildung 7.4: Mittelwerte des kognitiven Fragenniveaus pro Stunde⁷ mit veränderter Skala

Stunde	Einschätzung der Lehrkraft	Mittelwert des Fragenniveaus (4er-Skala)	Stunde	Einschätzung der Lehrkraft	Mittelwert des Fragenniveaus (4er-Skala)
GY 02	1	2.9	RS 01	0	1.3
GY 04	0	3.0	RS 03	0	2.3
GY 05	1	2.1	RS 06	2	2.0
GY 12	2	2.1	RS 07	1	1.5
GY 13	2	0.9	RS 08	0	1.1
GY 13b	2	4.0	RS 09	2	0.8
GY 14	1	2.7	RS 10	0	2.9
GY 15	4	1.9	RS 18	1	1.8
GY 16	1	2.3	RS 19	1	1.4
GY 17	1	4.0	RS 22	2	1.9
GY 20	0	1.6			
GY 21	2	0.0			

Tabelle 7.5: Vergleich zwischen der Einschätzung der Lehrkraft und dem Mittelwert des Fragenniveaus (N=22)⁸

⁷ Skala: ‚Reproduktionsfragen‘ (0), ‚Faktenfragen‘ (1), ‚Begründungsfragen‘ (2), ‚deep-reasoning-Fragen‘ (3)

Die Berechnung der Korrelation zeigt, dass kein systematischer Zusammenhang zwischen den Aussagen der Lehrkräfte im Interview zu Merkmalen der kognitiven Aktivierung im Klassengespräch und dem durchschnittlichen kognitiven Frageniveau in den jeweiligen Stunden besteht. Dies verdeutlicht das Streudiagramm.

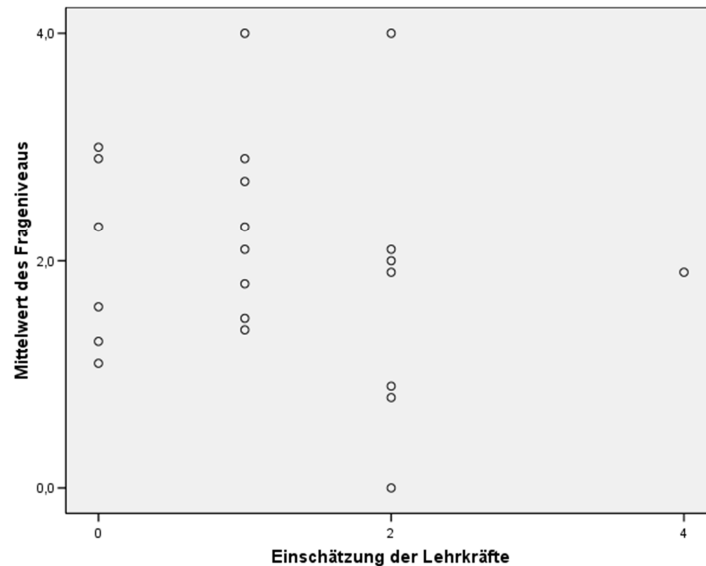


Abbildung 7.5: Streudiagramm zur Korrelation zwischen der Einschätzung der Lehrkräfte und dem Mittelwert des Frageniveaus

Der Zusammenhang zwischen genannten Merkmalen der kognitiven Aktivierung im Klassengespräch und dem mittleren kognitiven Frageniveau in der gefilmten Unterrichtsstunde der jeweiligen Lehrkraft interessiert insofern, als davon ausgegangen wurde, dass Aussagen zur kognitiven Aktivierung im Klassengespräch als handlungsleitende Kognitionen interpretiert werden können, die wiederum Einfluss auf das Frageverhalten der Lehrkräfte haben. Dieser Einfluss konnte innerhalb dieser Stichprobe nicht festgestellt werden.

⁸ Zu RS 11 existiert kein Mittelwert des kognitiven Frageniveaus, da in dieser Stunde das Klassengespräch ausschließlich zur Organisation der weiteren Projektarbeit genutzt wurde und keine inhaltlich relevanten Fragen gestellt wurden.

7.2.2 ZUSAMMENHANG ZWISCHEN VORBEREITUNG UND KOGNITIVEM NIVEAU DER FRAGE

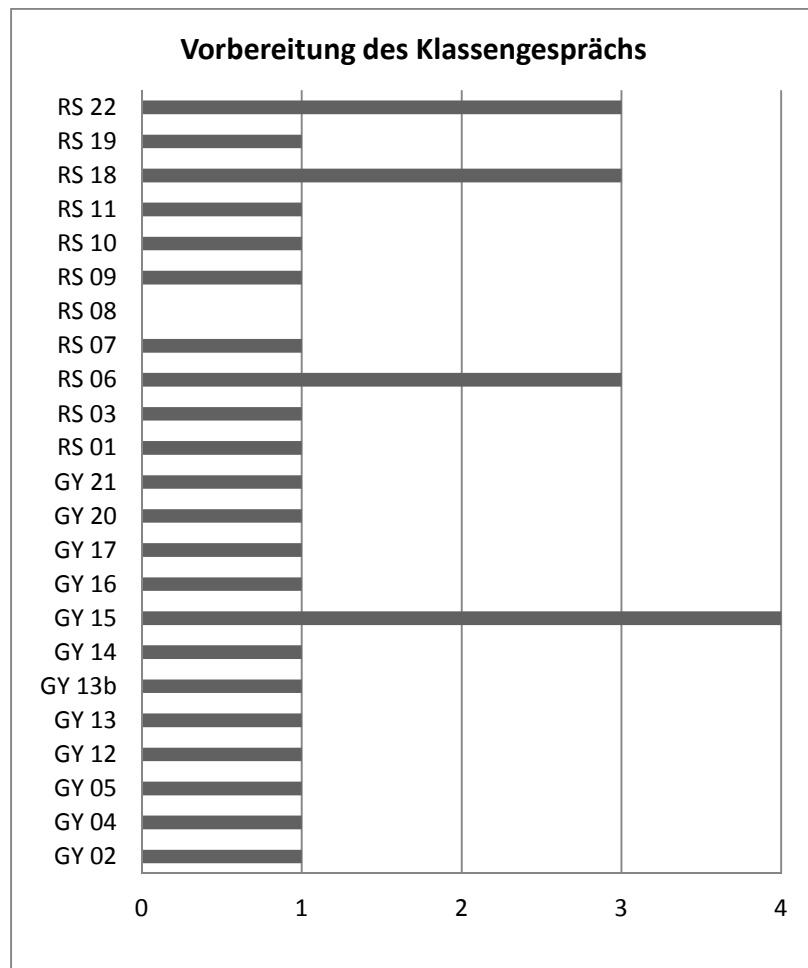


Abbildung 7.6: Vorbereitung des Klassengesprächs aus Sicht der Lehrkräfte.

Die Lehrkräfte wurden in den Interviews auch gezielt zur Vorbereitung im Hinblick auf das Klassengespräch gefragt. Bei der Auswertung des Zusammenhangs zwischen den Aussagen der Lehrkräfte zur Vorbereitung und dem mittleren kognitiven Niveau der Fragen im Klassengespräch wurde analog verfahren wie in Abschnitt 7.2.1 beschrieben.

Die Skala ist in diesem Fall so zu interpretieren, dass der Wert 0 die ausdrückliche Ablehnung von Vorbereitung bedeutet. Lehrkräfte, deren Aussagen mit 1 kodiert sind, äußerten sich entweder gar nicht zur Vorbereitung des Klassengesprächs oder eher ablehnend, ebenfalls mit der Begründung, dass die Vorbereitung sie einschränke. Somit legen nur vier Lehrkräfte Wert auf die Vorbereitung von Fragen und Impulsen oder die Vorstrukturierung des Gesprächs und machen sich zu Hause bei der Unterrichtsvorbereitung Gedanken über die Gesprächsphasen. Die Irritation über die Frage nach der Vorbereitung des Klassengesprächs

bei einigen Lehrkräften während des Interviews kann zusätzlich dahingehend interpretiert werden, dass Klassengespräche nie (bewusst) vorbereitet werden. Die Frage zu den vorbereitenden Gedanken in Bezug auf die Aufgabenwahl am Anfang des Interviews führte hingegen zu keinen Irritationen.

Stunde	Einschätzung der Lehrkraft	Mittelwert des Frageniveaus (4er-Skala)	Stunde	Einschätzung der Lehrkraft	Mittelwert des Frageniveaus (4er-Skala)
GY 02	1	2.9	RS 01	1	1.3
GY 04	1	3.0	RS 03	1	2.3
GY 05	1	2.1	RS 06	3	2.0
GY 12	1	2.1	RS 07	1	1.5
GY 13	1	0.9	RS 08	0	1.1
GY 13b	1	4.0	RS 09	1	0.8
GY 14	1	2.7	RS 10	1	2.9
GY 15	4	1.9	RS 18	3	1.8
GY 16	1	2.3	RS 19	1	1.4
GY 17	1	4.0	RS 22	3	1.9
GY 20	1	1.6			
GY 21	1	0.0			

Tabelle 7.6: Vergleich zwischen der Vorbereitung und dem Mittelwert des Frageniveaus (N=22)

Wie die Verteilung der einzelnen Daten bereits nahe legt, besteht kein Zusammenhang zwischen den Aussagen der Lehrkraft zur Vorbereitung des Klassengesprächs und dem mittleren kognitiven Frageniveau der gefilmten Stunde. Dies könnte darauf zurückzuführen sein, dass die Vorbereitung sich vor allem auf die Strukturierung des Gesprächs bezieht und weniger darauf, kognitiv anspruchsvolle Fragen und Impulse vorzubereiten. Desweiteren können im Unterschied zur Vorbereitung von Aufgaben nicht alle auftretenden Situationen im Klassengespräch bei der Vorbereitung vorausgedacht werden, so dass die Lehrkraft in der konkreten Unterrichtssituation darauf angewiesen ist, spontan zu reagieren. Auch diese unvorbereiteten Fragen fließen in die Mittelwertberechnung ein.

8 DISKUSSION

Im abschließenden Kapitel werden die Ergebnisse aus dem vorherigen Kapitel im Hinblick auf die gestellten Forschungsfragen zusammengefasst. Zudem wird auf methodische Schwierigkeiten verwiesen und aufgezeigt, wie die Forschung zum Thema Fragen und Impulse im Klassengespräch fortgeführt werden könnte.

8.1 ZUSAMMENFASSUNG DER ZENTRALEN BEFUNDE

Hinsichtlich der Frage, ob das durchschnittliche Frageniveau schulartspezifisch zu unterscheiden ist, kann festgehalten werden, dass die Überlegungen, die sich aus der Literatur ergaben, mit den Forschungsergebnissen der vorliegenden Arbeit übereinstimmen. Für das Gymnasium konnte ein signifikant höherer Mittelwert des kognitiven Frageniveaus ermittelt werden. Die kognitiv akzentuierte Schulkultur, die dem Gymnasium zugeschrieben wird, schlägt sich auch im Frageniveau nieder. Nochmals bemerkt wird an dieser Stelle, dass kognitive Aktivierung nicht mit dem erhöhten Schwierigkeitsgrad einer Frage gleichzusetzen und daher in gleicher Weise an der Hauptschule möglich ist.

Die Analyse der gestellten Fragen ohne Berücksichtigung von deren Beantwortung zeigte, dass auch an der Real- und Hauptschule kognitiv aktivierende Fragen gestellt werden. Gerade in der Häufigkeit der Begründungsfragen, die auch noch reproduktiven Charakter haben, aber auf komplexere Antworten der Schülerinnen und Schüler zielen, unterscheidet sich das Gymnasium von den beiden anderen Schularten. Das Problem gerade an der Hauptschule scheint darin zu liegen, dass viele kognitiv anspruchsvolle Fragen nicht so bearbeitet werden können, wie die Lehrkraft dies intendiert. Die Reaktion der Lehrkraft in solchen Fällen wurde im Rahmen dieser Arbeit nicht untersucht.

Die Anzahl der pro Minute gestellten Fragen ist an der Hauptschule deutlich höher als am Gymnasium. Hauptschullehrkräfte stellen pro Minute etwa 1,5-mal so viele Fragen wie Gymnasiallehrkräfte. Auch dies deutet auf ein geringeres kognitives Niveau der Fragen hin, denn im Rahmen anderer Untersuchungen konnte belegt werden, dass anspruchsvolle Fragen einer längeren Antwortzeit bedürfen. Dies konnte auch im Rahmen dieser Arbeit bestätigt werden: Zum einen korreliert eine steigende Fragefrequenz sehr hoch mit der

Abnahme des Fragenniveaus, zum anderen weisen Unterrichtsstunden, in denen die Schülerinnen und Schüler weniger Redeanteile am Klassengespräch erhalten, ein niedriges Fragenniveau auf. Beide Ergebnisse zusammen weisen darauf hin, dass die Schülerinnen und Schüler am Gymnasium nicht nur stärker in das Klassengespräch einbezogen werden, sondern auch, dass an gestellten Fragen und Impulsen im Gymnasialunterricht länger gearbeitet wird.

Die Ergebnisse, die aus der niedrig-inferenten Basiskodierung und der Analyse des Fragenniveaus gewonnen wurden, konnten durch die Analyse einzelner Items zur kognitiven Aktivierung im Klassengespräch aus der hoch-inferenten Analyse bestätigt werden. Bei drei Items zur Einschätzung der kognitiven Aktivierung wiesen die Gymnasialstunden im Mittel den größten Wert auf. In Bezug auf die Offenheit der Klassengespräche kann davon ausgegangen werden, dass dies in den Klassengesprächen am Gymnasium größtenteils realisiert ist. Unterstützt wird dies durch den niedrig-inferenten Befund zur Fragefrequenz und zum Redeanteil der Schülerinnen und Schüler. Auffallend ist, dass das Vorwissen an der Hauptschule am stärksten aktiviert wird. Auch dies lässt sich mit den Befunden der niedrig-inferenten Analyse in Einklang bringen, wenn das Vorwissen vorwiegend über Reproduktions- und Faktenfragen abgefragt und nicht zur Erklärung neuer Sachverhalte herangezogen wurde. An dieser Stelle muss jedoch darauf hingewiesen werden, dass das Item zur Aktivierung des Vorwissens eine sehr geringe Beurteilerübereinstimmung aufweist und die Interpretation der Ergebnisse insofern eingeschränkt ist.

Die Realschule nimmt bei allen Analysen eine Mittelposition zwischen Gymnasium und Hauptschule ein. Dabei liegen die Mittelwerte meistens näher an denen der Hauptschule.

Die Interviewaussagen der Lehrkräfte wurden in zweierlei Hinsicht ausgewertet: Zum einen wurde bewertet, wie viel Wert Lehrkräfte nach eigener Aussage auf Merkmale der kognitiven Aktivierung im Klassengespräch legen wie etwa auf die Offenheit des Gesprächs, auf den Bezug der Schülerinnen und Schüler untereinander in ihren Beiträgen, auf die Beteiligung möglichst vieler Schülerinnen und Schüler oder auf die Ausführlichkeit der Schülerbeiträge. Der Vergleich mit dem durchschnittlichen kognitiven Fragenniveau der jeweiligen Stunde ergab keinen Zusammenhang zwischen dem Wert, den die Lehrkraft nach

eigener Aussage auf Merkmale der kognitiven Aktivierung legt, und dem Frageniveau. Hier müsste untersucht werden, inwiefern sich handlungsleitende Kognitionen zu einzelnen Merkmalen der kognitiven Aktivierung auf deren Realisierung im Unterricht auswirken, also ob Lehrkräfte, die beispielsweise nach eigener Aussage Wert auf Beteiligung möglichst vieler Schülerinnen und Schüler legen, ein entsprechendes Aufrufverhalten zeigen.

Im Hinblick auf die Lehreraus- und –fortbildung erscheint interessant, inwiefern sich die Vorbereitung des Klassengesprächs auf das durchschnittliche Frageniveau der Stunde auswirkt. Auch hier konnte kein Zusammenhang belegt werden. Bestätigt hat sich aber die Vermutung, dass Lehrkräfte wenig bis gar keine Zeit für die Vorbereitung der Gesprächsphasen aufwenden.

Das hauptsächliche Forschungsinteresse dieser Arbeit lag darin, schulartspezifisch Fragen und Impulse im Klassengespräch zu untersuchen. Es lässt sich abschließend feststellen, dass sich deutliche Unterschiede zwischen den Schularten unter verschiedenen Fragestellungen in Bezug auf das Klassengespräch feststellen lassen. Zusammenfassend kann festgehalten werden, dass in Bezug auf die kognitive Aktivierung das Klassengespräch am Gymnasium stets die höchsten Werte aufweist.

8.2 METHODISCHE ÜBERLEGUNGEN UND FOLGERUNGEN FÜR WEITERFÜHRENDE AUSWERTUNGEN

Abschließend muss darauf hingewiesen werden, dass viele der oben beschriebenen Ergebnisse auf Daten beruhen, die nicht die notwendige Reliabilität aufweisen. Die geringen Übereinstimmungswerte beruhen unter anderem darauf, dass im Rahmen dieser Abschlussarbeit mit verschiedenen Methoden exemplarisch gearbeitet wurde. Dies hatte den Vorteil, dass die Autorin ihre Methodenkompetenz erweitern konnte. Nachteilig wirkte sich dies jedoch auf die Auswertung der Daten aus. So wurde beispielsweise bei der Entwicklung des Kodiermanuals für die Auswertung des kognitiven Niveaus der Fragen der zyklische Prozess der quantifizierenden Inhaltsanalyse nicht eingehalten und Probekodierungen mit anschließender Überarbeitung des Manuals nicht durchgeführt. Zudem fand die Zweitkodierung der Fragen ohne Kodiererschulung im Vorfeld statt. Die Weiterentwicklung des Videoanalysemanuals um die Items zum Klassengespräch wurde ebenfalls nicht durch Probekodierungen überprüft. Der Ursache für die geringen

Übereinstimmungswerte bei der Basiskodierung „Art der Lehreräußerungen im Klassengespräch“ wurde aus Zeitgründen nicht nachgegangen. Die Ergänzung des Interviewleitfadens um Fragen zum Klassengespräch musste zu einem Zeitpunkt erfolgen, als die Zielsetzung dieser Arbeit nur ansatzweise feststand. Auf der anderen Seite ermöglichte die Mitarbeit in einem Forscherteam neben dem Ausbau der methodischen Fertigkeiten vor allem eine höhere Konstruktvalidität.

Gute Übereinstimmungswerte zeigten sich stets dann, wenn beide Kodierer bzw. Rater intensiv an der Erarbeitung des jeweiligen Manuals beteiligt waren. Dies ist ein methodischer Aspekt, der für die weitere Auswertungen zu berücksichtigen ist. Bedenkenswert ist ebenfalls, ob die Bewertung aller Gesprächsphasen einer Unterrichtsstunde mit jeweils einem Item nicht zu Verzerrungen der Ergebnisse geführt hat und ob es daher nicht sinnvoller wäre, analog zur Auswertung der Phasen in Bezug auf die Aufgabenbearbeitung, jede Gesprächsphase einzeln zu bewerten und Mittelwerte zu bilden.

Nicht berücksichtigt wurden in dieser Arbeit beispielsweise die Einzelverläufe von Unterrichtsgesprächen. KUNTZE & REISS (2004) identifizierten Figuren der Fragetechnik:

- Schwer-einfach-Zyklen: Eine anspruchsvolle Ausgangsfrage (Kategorie 3 oder 4) bleibt unbeantwortet oder wird nicht zur Zufriedenheit der Lehrkraft beantwortet und daher über Nachfragen kleingearbeitet (Kategorie 1 oder 2).
- Frage-Insistierung: Eine Frage ohne oder mit unvollständiger bzw. falscher Antwort wird erneut (im wesentlichen Wortlaut ohne weitere Hilfestellung) gestellt.

Solche Fragefiguren könnten in einem weiteren Analyseschritt identifiziert werden. Die im Rahmen dieser Arbeit entstandenen Instrumente könnten nach einer Überarbeitung und erneuten Reliabilitätsprüfung dafür genutzt werden.

9 LITERATUR

- BLOOM, BENJAMIN S. (HRSG.) (1972). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim u.a.: Beltz.
- BOHL, THORSTEN (2003). Anstöße – PISA und Lernende in der Hauptschule. In: BOHL, THORSTEN ET AL. (HRSG.). *Lernende in der Hauptschule: ein Blick auf die Hauptschule nach PISA*. 1. Aufl., Baltmannsweiler: Schneider Verlag Hohengehren, S. 15-63.
- BOHL, THORSTEN ET AL. (im Druck). *Aufgabenkultur in verschiedenen Schularten: eine vergleichende Analyse von Aufgaben und Lehrerhandeln im Hauptschul-, Realschul- und Gymnasialunterricht*. Abschlussbericht zum DFG-Projekt (BO 3312/1-1).
- BOS, WILFRIED & TARNAI, CHRISTIAN (1999). Content analysis in empirical social research. *International Journal of Educational Research* 31 (8), S. 659-671.
- CLAUSEN, MARTEN ET AL. (2003). Unterrichtsqualität auf der Basis hoch-inferenter Unterrichtsbeurteilungen. *Unterrichtswissenschaft* 31(2), S. 122-141.
- DUBBERKE, THAMAR ET AL. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. *Zeitschrift für Pädagogische Psychologie* 22 (3-4), S. 193-206.
- EINSIEDLER, WOLFGANG (2002). Das Konzept „Unterrichtsqualität“. *Unterrichtswissenschaft* 30 (3), S. 194-196.
- GLÄSER, EVA (2004). Mehr als nur darüber reden. *Grundschule* 36 (5), S. 36-38.
- GROEBEN, ANNEMARIE VON DER (2009). Wie viel Führung braucht das Unterrichtsgespräch? *Pädagogik* 61 (1), S. 6-10.
- HELFERICH, CORNELIA (2004). *Die Qualität qualitative Daten*. Manual für die Durchführung von Interviews. Wiesbaden: VS-Verlag.
- HELMKE, ANDREAS (2007). *Unterrichtsqualität: erfassen, bewerten, verbessern*. 6. Aufl., Seelze: Kallmeyer; Stuttgart: Klett.

- HUGENER, ISABELLE (2006). Sozialformen und Lektionsdauer. In: HUGENER, ISABELLE ET AL. (HRSG.). Videoanalysen. Aus: KLIEME, ECKHARD ET AL. (HRSG.). *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“*, Teil 3. Materialien zur Bildungsforschung, Band 15. Frankfurt a. M.: GFPPF, S. 55-61.
- HUGENER, ISABELLE ET AL. (2006). Videobasierte Unterrichtsforschung: Integration verschiedener Methoden der Videoanalyse für eine differenzierte Sicht auf Lehr-Lernprozesse. In: SYBILLE RAHM ET AL. (HRSG.). *Unterrichtsforschung - Perspektiven innovativer Ansätze*. Innsbruck, Wien, Bozen: Studien-Verlag, S. 41-53.
- HUGENER, ISABELLE (2008). *Inszenierungsmuster im Unterricht und Lernqualität: Sichtstrukturen schweizerischen und deutschen Mathematikunterrichts in ihrer Beziehung zu Schülerwahrnehmung und Lernleistung - eine Videoanalyse*. Münster, München, Berlin u.a.: Waxmann.
- JATZWAUK, PAUL (2007). *Aufgaben im Biologieunterricht – eine Analyse der Merkmale und des didaktisch-methodischen Einsatzes von Aufgaben im Biologieunterricht*. Berlin: Logos.
- JATZWAUK, PAUL ET AL. (2008). Der Einfluss des Aufgabeneinsatzes im Biologieunterricht auf die Lernleistung der Schüler – Ergebnisse einer Videostudie. *ZfDN* 14, S. 263-283.
- KIPER, HANNA (2005). Pädagogik des Gymnasiums – quo vadis? *PÄD-Forum* 33 (5), S. 301-306.
- KLEINKNECHT, MARC (2010). *Aufgabenkultur im Unterricht: eine empirisch-didaktische Video- und Interviewstudie an Hauptschulen*. Baltmannsweiler: Schneider Verl. Hohengehren.
- KLIEME, ECKHARD (2002). Was ist guter Unterricht? In: BERGSDORF, WOLFGANG (HRSG.). *Herausforderungen der Bildungsgesellschaft*. Weimar: RhinoVerlag, S. 89-113.
- KLIEME, ECKHARD (2006). Empirische Unterrichtsforschung: aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. *Z.f.Päd* 52 (6), S. 765-773.
- KNIGGE, MICHEL (2009). *Hauptschüler als Bildungsverlierer?: eine Studie zu Stigma und selbstbezogenem Wissen bei einer gesellschaftlichen Problemgruppe*. Münster, München, Berlin u.a.: Waxmann.

- KOBARG, MAREIKE & SEIDEL, TINA (2003). Prozessorientierte Lernbegleitung im Physikunterricht. In: SEIDEL, TINA ET AL. (HRSG.). *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“*. Kiel: IPN, S. 151-200.
- KUNTZE, SEBASTIAN & REISS, KRISTINA (2004). Unterschiede zwischen Klassen hinsichtlich inhaltlicher Elemente und Anforderungsniveaus im Unterrichtsgespräch beim Erarbeiten von Beweisen. Ergebnisse einer Videoanalyse. *Unterrichtswissenschaft* 32 (4), S. 357-379.
- LEUDERS, TIMO (2007). Fachdidaktik und Unterrichtsqualität im Bereich Mathematik. In: ARNOLD, KARL-HEINZ (HRSG.). *Unterrichtsqualität und Fachdidaktik*. Bad Heilbrunn: Klinkhardt.
- LIPOWSKY, FRANK ET AL. (2005). Unterrichtsqualität im Schnittpunkt unterschiedlicher Perspektiven. In: HOLTAPPELS, HEINZ GÜNTHER ET AL. (HRSG.). *Schulentwicklung und Schulwirksamkeit: Systemsteuerung, Bildungschancen und Entwicklung der Schule*. Weinheim, München: Juventa, S. 223-238.
- LIPOWSKY, FRANK ET AL. (2008). Schülerbeteiligung und Unterrichtsqualität. In: GLÄSER-ZIKUDA, MICHAELA ET AL. (HRSG.). *Lehrerexpertise: Analyse und Bedeutung unterrichtlichen Handelns*. Münster, München, Berlin u.a.: Waxmann, S. 67-90.
- LOWYCK, JOOST (1976). Die Analyse des Fragenstellens als Instrument für ein abgestuftes Fertigkeitstraining. *Unterrichtswissenschaft* 4(1), S. 53-73.
- MAYRING, PHILIPP (2008). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 10., neu ausgestattete Aufl., Weinheim, Basel: Beltz.
- MÜLLER, CHRISTOPH THOMAS (2004). *Subjektive Theorien und handlungsleitende Kognitionen von Lehrern als Determinanten schulischer Lehr-Lern-Prozesse im Physikunterricht*. Berlin: Logos-Verl.
- MÜLLER, CHRISTIANE ET AL. (2006). Chancen und Grenzen von Videostudien in der Unterrichtsforschung. In: SYBILLE RAHM ET AL. (HRSG.). *Unterrichtsforschung - Perspektiven innovativer Ansätze*. Innsbruck, Wien, Bozen: Studien-Verlag, S. 125-138.

- NIEGEMANN, HELMUT & STADLER, SILKE (2001). Hat noch jemand eine Frage? Systematische Unterrichtsbeobachtung zu Häufigkeit und kognitivem Niveau im Unterricht. *Unterrichtswissenschaft* 29 (2), S. 171-192.
- ORTH, PETER (2009). Gesprächsformen und ihre Gelingensbedingungen. *Pädagogik* 61 (1), S. 30-33.
- PAULI, CHRISTINE (2006). Das Klassengespräch. In: HUGENER, ISABELLE ET AL. (HRSG.). Videoanalysen. Aus: KLIEME, ECKHARD ET AL. (HRSG.). *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“*, Teil 3. Materialien zur Bildungsforschung, Band 15. Frankfurt a. M.: GFPPF, S. 124-147.
- PAULI, CHRISTINE & LIPOWSKY, FRANK (2007). Mitmachen oder zuhören? Mündliche Schülerinnen und Schülerbeteiligung im Mathematikunterricht. *Unterrichtswissenschaft* 35 (2), S. 101-124.
- PAULI, CHRISTINE ET AL. (2008). Kognitive Aktivierung im Mathematikunterricht. *Zeitschrift für Pädagogische Psychologie* 22(2), S. 127-133.
- REYER, THOMAS (2004). *Oberflächenmerkmale und Tiefenstrukturen im Unterricht: exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe*. Berlin: Logos-Verl.
- SEIDEL, TINA (2003a). *Lehr-Lernskripts im Unterricht: Freiräume und Einschränkungen für kognitive und motivationale Lernprozesse - eine Videostudie im Physikunterricht*. Münster, München, Berlin u.a.: Waxmann.
- SEIDEL, TINA (2003b). Videobasierte Kodierverfahren in der IPN Videostudie Physik – ein methodischer Überblick. In: SEIDEL, TINA ET AL. (HRSG.). *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“*. Kiel: IPN, S. 99-112.
- SEIDEL, TINA (2003c). Sichtstrukturen – Organisation unterrichtlicher Aktivitäten. In: SEIDEL, TINA ET AL. (HRSG.). *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“*. Kiel: IPN, S. 113-128.

- SEIDEL, TINA ET AL. (2003a). Gelegenheitsstrukturen beim Klassengespräch und ihre Bedeutung für die Lernmotivation. *Unterrichtswissenschaft* 31 (2), S. 142-165.
- SEIDEL, TINA ET AL. (2003b). Aufzeichnen von Physikunterricht. In: SEIDEL, TINA ET AL. (HRSG.). *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“*. Kiel: IPN, S. 47-76.
- SEIDEL, TINA & PRENZEL, MANFRED (2004). Muster unterrichtlicher Aktivität im Physikunterricht. In: DOLL, JÖRG ET AL. (HRSG.). *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung*. Münster, München, Berlin u.a.: Waxmann, S. 177-194.
- SEIDEL, TINA (im Druck). Lehrerhandeln im Unterricht. In: TERHART, EWALD ET AL. (Hrsg.). *Handbuch der Forschung zum Lehrerberuf*. Münster: Waxmann, S. 549-573.
- SMAWIL, JULIA (2008). *Gewinner und Verlierer des dreigliedrigen Schulsystems in NRW bei Schülern und Lehrern: Bildungsprozesse aus psychologischer Perspektive*. Duisburg, Essen, Univ., Diss.
- TRAUTWEIN, ULRICH ET AL. (2007). Hauptschulen = Problemschulen? *APuZ* 28, S. 3-9.
- VYGOTSKIJ, LEV S. (1979). *Mind in society: the development of higher psychological processes*. 2. Aufl., Cambridge, Mass.: Harvard Univ. Pr.

10 EIDESSTATTLICHE ERKLÄRUNG

Batzel, Andrea

3151202

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit bzw. Studienleistung zum Thema

Fragen und Impulse im Klassengespräch

eigenständig, ohne unerlaubte Hilfe und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe. Alle sinngemäß und wörtlich übernommenen Textstellen aus Veröffentlichungen oder aus anderweitigen fremden Äußerungen habe ich als solche einzeln kenntlich gemacht.

Mir ist bekannt, dass ich im Falle einer Täuschung das Recht auf den Erwerb der Master-Prüfung im Studiengang

Forschung und Entwicklung in der Erziehungswissenschaft

verwirkt habe.

Weiterhin erkläre ich, dass ich die vorliegende Arbeit bzw. Studienleistung in dieser oder in leicht veränderter Form in keiner anderen Lehrveranstaltung zum Zwecke des Erwerbs eines Leistungsnachweises eingereicht habe.

Tübingen, den 11. August 2010

11 ANHANG

Im Anhang werden nur die in eigenständiger Leistung entwickelten Erhebungs- und Auswertungsinstrumente aufgeführt. Die weiteren Instrumente finden sich in BOHLE ET AL. (im Druck).

- Übersichtstabelle mit Daten aus der niedrig-inferenten Basiskodierung
- Kodiermanual für das Einschätzen des kognitiven Niveaus der Fragen und Impulse im Klassengespräch
- Auswertungstabellen zur Einschätzung des kognitiven Niveaus der Fragen zu den einzelnen Stunden