# II Anlage des gesamten Forschungsprojekts

Teil II gibt einen Überblick über das Dissertationsprojekt und seine Einbettung in den übergeordneten Forschungskontext (Kapitel 8). Zunächst finden sich Ausführungen zum Mixed-Methods-Design im Allgemeinen (Abschnitt 7.1) und in der konkreten Ausführung (Abschnitt 7.2). Abschließend werden der zeitliche Ablauf und die Phasen des Forschungsprojekts dargelegt (Kapitel 9).

# 7 Quantitative und qualitative Verfahren und ihre Triangulation: Mixed-Methods-Design

Während die Breitenstudie durch quantitative Verfahren geprägt ist (vgl. Kapitel 10), umfassen die Tiefenstudien sowohl quantitative, als auch hauptsächlich qualitative Anteile. Zur Ermöglichung von Transparenz über die Frage, welche Forschungsansätze aus welchem Grund herangezogen wurden, werden im Folgenden einige allgemein-methodologische Überlegungen für quantitative und qualitative Verfahren und ihre Triangulation angeführt. Darauf aufbauend wird dargelegt, wie die Triangulation im Forschungsprozess abgelaufen ist.

## 7.1 Quantitative und qualitative Forschungsansätze und ihre Triangulation

Die Unterscheidung quantitativer und qualitativer Verfahren kann sowohl hinsichtlich der Datenerhebung, als auch der Datenauswertung vorgenommen werden. Quantitative Forschungsansätze zeichnen sich – beide Phasen betreffend – durch einen relativ linearen Ablauf aus: So wird eine Forschungsfrage zunächst präzisiert und strukturiert und anschließend durch ein Instrument zur Datenerhebung operationalisiert. Die Daten werden nach der Erhebung gemäß einer festgelegten Systematik ausgewertet (Burzan 2005, S. 25). Die Wissenschaftlichkeit quantitativer Forschung bemisst sich anhand der Gütekriterien Objektivität, Reliabilität und Validität (Hussy, Schreier & Echterhoff 2013, S. 23f).

Im Gegensatz dazu wird in qualitativer Forschung der Objektivitätsanspruch ersetzt durch intersubjektive Nachvollziehbarkeit (zu diesem und weiteren Gütekriterien vgl. Steinke 2000). Der Mittelpunkt der Forschungsbemühungen von Vertreterinnen und Vertretern der qualitativen Tradition kann folgendermaßen beschrieben werden:

> „Erkundung der Sinnsetzungs- und Sinndeutungsvorgänge der Akteure im Untersuchungsfeld, die Exploration kultureller Praktiken und Regeln und die genaue und tiefgehende Analyse und Beschreibung von Einzelfällen." (Kelle 2008, S. 13)

© Springer Fachmedien Wiesbaden 2016
N. Wilhelm, *Zusammenhänge zwischen Sprachkompetenz und Bearbeitung mathematischer Textaufgaben*, Dortmunder Beiträge zur Entwicklung und Erforschung des Mathematikunterrichts 25, DOI 10.1007/978-3-658-13736-6_2

Zur Erreichung dieser Zielsetzung mit dem Fokus auf die intensivere Untersuchung von einzelnen Fällen kommen beim qualitativen Ansatz „eher sinnverstehende (unstandardisierte) Verfahren zum Einsatz“ (Hussy et al. 2013, S. 9).

Lange Zeit standen quantitative und qualitative Verfahren in Konkurrenz zueinander und das Verhältnis wurde durch wechselseitige Abgrenzung und Kritik gekennzeichnet (zu näheren Ausführungen vgl. u.a. Kelle 2008, S. 25ff). Mittlerweile wird eher weniger über die Überlegenheit eines Forschungsansatzes diskutiert, sondern die Auswahl von der Adäquatheit abhängig gemacht. Diese richtet sich nach dem Forschungsgegenstand und dem Forschungsinteresse (Hussy et al. 2013, S. 10). Zunehmend werden auch beide Ansätze aufeinander bezogen.

Für den empirischen Teil der vorliegenden Arbeit ist die Triangulation von quantitativen und qualitativen Methoden zentral. Methodisch bedeutet dies eine wechselseitige Ergänzung, „wobei diese […] in der komplementären Kompensation der Schwächen und blinden Flecke der jeweiligen Einzelmethode liegt“ (Flick 2009, S. 44). Darauf wird im folgenden Abschnitt 7.2 konkret hinsichtlich der realisierten Studie eingegangen.

Die Kombination quantitativer und qualitativer Methoden ist in der Forschungspraxis unter einer Vielzahl von Bezeichnungen und Ansätzen etabliert: Neben dem in dieser Arbeit verwendeten Terminus und Ansatz *Mixed-Methods* ist insbesondere im englischsprachigen Raum auch von *integrative research, mixed research, multimethods* usw. die Rede mit oftmals unterschiedlicher inhaltlicher Ausgestaltung (für eine ausführliche Darlegung der Ansätze vgl. Creswell & Plano Clark 2007). Die Unterscheidung der Ansätze verläuft hauptsächlich über die Festlegung, in welcher Gewichtung und Reihenfolge quantitative und qualitative Elemente eingesetzt werden. Folgende Typen von Designs, die qualitative und quantitative Forschung integrieren, sind möglich (Flick 2009, S. 42ff; ähnlich in Tashakkori & Teddlie 1998):

- Qualitative und quantitative Elemente finden parallel statt.
- Die kontinuierliche Feldbeobachtung liefert die Basis, verschiedene Erhebungswellen von Umfragen beziehen sich darauf.
- Eine qualitative Erhebung stellt den Beginn dar, eine Fragebogenstudie schließt sich an. Abschließend werden die Ergebnisse aus beiden Schritten in einer zweiten qualitativen Phase vertieft und überprüft.
- Eine Umfrage stellt die Basis dar. Daran schließt sich eine Feldstudie an, welche die Umfrageergebnisse ergänzt und vertieft. Durch eine experimentelle Intervention werden die Ergebnisse der ersten beiden Schritte überprüft.

Wichtiger als diese formale Ebene der genauen Einhaltung eines festgeschriebenen Verlaufs ist in der vorliegenden Arbeit die Offenlegung, welche quantitativen und qualitativen Elemente zu welchem Zweck eingesetzt werden

und welche Rolle sie im Forschungsprozess spielen. Genauer steht die Beantwortung folgender Fragen im Vordergrund:

> „Für welche […] Untersuchungsgegenstände sind eher qualitative, für welche Forschungsfragen eher quantitative Verfahren geeignet? Welche Methodenprobleme qualitativer und quantitativer Forschung werden durch bestimmte Untersuchungsgegenstände erzeugt? Wie können Stärken beider Methodentraditionen genutzt werden, um diese Methodenprobleme zu lösen?“ (Kelle 2008, S. 15)

Kelle (2008, S. 233ff) liefert eine Zusammenstellung der Möglichkeiten von Methodenkombinationen in Mixed-Methods-Designs:

- *Erklärung überraschender statistischer Befunde*: Bei einem statistischen Zusammenhang können spezifische erklärende Variablen die Varianz der abhängigen Variable meistens nicht oder nur bis zu einem gewissen Grad erklären. Die Erklärung ist aber immer mit Interpretationsschwierigkeiten verbunden, da immer nur einzelne Elemente aus einer Fülle von tatsächlich wirkenden betrachtet werden. Welche Rolle sie tatsächlich bei der Erklärung spielen, kann mit qualitativen Zugängen, z.B. Interviews, aufgedeckt werden.
- *Identifikation von Variablen, die bislang unerklärte statistische Varianz aufklären*: Weitere erklärende Elemente, die bei einer statistischen Erhebung nicht sichtbar werden, können erst mit qualitativen Zugängen, aufgedeckt werden und somit Erklärungsansätze bieten.
- *Untersuchung der Geltungsreichweite von qualitativ entwickelten Kategorien*: Mit Hilfe von qualitativen Methoden können zwar kausale Pfade entdeckt werden, allerdings aufgrund forschungspragmatischer Beschränkungen immer nur für eine kleine Anzahl von Fällen. Quantitative Methoden geben hier die Möglichkeit, Grenzen der Verallgemeinerbarkeit überwinden zu können.
- *Steuerung der Fallauswahl in qualitativen Studien*: Methodenkombination kann nicht nur im Nachhinein zur Verallgemeinerbarkeit qualitativer Ergebnisse eingesetzt werden, sondern auch „um die Verallgemeinerungsfähigkeit der Daten und Befunde von vornherein zu sichern“ (Kelle 2008, S. 247). Mit Hilfe statistischer Verfahren findet eine bewusste Stichprobenziehung statt, indem der Fokus auf bestimmte Merkmalsausprägungen gelegt wird, um z.B. Fälle kontrastieren zu können.
- *Aufdeckung und Beschreibung von Methodenartefakten in qualitativen und quantitativen Studien*: Die Kombination quantitativer und qualitativer Verfahren ermöglicht, Methodenartefakte aufzudecken, die sich in den meisten Fällen auf die Gültigkeit und Zuverlässigkeit der Ergebnisse beziehen. So können z.B. Wissenslücken bei Fragebögen zu ungültigen Ergebnissen führen. Im umgekehrten Fall des Einsatzes von qualitativen Verfahren können Einflussfaktoren z.B. durch die Steuerung von Interviewteilnehmenden

verborgen bleiben, die durch statistische Verfahren aufgedeckt werden können.

Die Auflistung macht deutlich, dass weder qualitative oder quantitative Zugänge als überlegen gelten können. Sie werden hier vielmehr als sich gegenseitig ergänzend gesehen. Die Rolle der jeweiligen Ansätze in der vorliegenden Arbeit wird im nachfolgenden Abschnitt 7.2 detaillierter im Hinblick auf die Phasen des Dissertationsprojekts beschrieben.

## 7.2 Mixed-Methods-Design in der vorliegenden Arbeit

Im Folgenden sollen die bisher dargestellten Aspekte zu Mixed-Methods-Designs auf die vorliegende Arbeit angewendet werden. Dabei stellt sich zunächst die Frage nach der Gewichtung von quantitativen und qualitativen Elementen. Auch wenn qualitativen Zugängen in der Tiefenstudie II (vgl. Teil V) ein bedeutend größerer Anteil eingeräumt wird als quantitativen Elementen, werden sie einander von ihrer Aussagekraft her nicht pauschal übergeordnet. Vielmehr wurden sie gemäß ihrer Passung zum Untersuchungsinteresse ausgewählt: Qualitative Verfahren wie erste Interviews dienten zunächst einer Sensibilisierung für den Forschungsgegenstand im Allgemeinen. Es wurde dabei vordergründig das Ziel verfolgt, Einblicke in Bearbeitungsprozesse zu gewinnen, um beobachtbare Schwierigkeiten erfassen zu können. Quantitative Vorgehensweisen wie die Analyse der vorliegenden Items gaben Anhaltspunkte dafür, welche Items potenzielle Hürden enthalten und welche Merkmale ggf. Schwierigkeiten erzeugen. Qualitative Vorgehen in Form weiterer Interviews ermöglichten dann wiederum, mögliche Erklärungsansätze für statistisch festgestellte Zusammenhänge zu ermitteln. Konkret bedeutet dies, dass durch Interviews Schwierigkeiten im Lösungsprozess bei jenen Items analysiert wurden, die sich in den DIF-Analysen als schwierig für sprachlich Schwache herausstellten (Erklärung überraschender statistischer Befunde und Identifikation von Variablen, die bislang unerklärte statistische Varianz aufklären, vgl. Abschnitt 7.1).

Zusätzlich wurden in der ersten Interviewserie auch Items in den Blick genommen, die nicht auffällig bzgl. ihres DIF-Werts waren, aber in der Bearbeitung spezifische Vorstellungen und Schwierigkeiten von sprachlich schwachen Lernenden zeigten. In Teilbereichen war es auf Grundlage der Interviewstudie wiederum möglich, Sachverhalte quantitativ zu untersuchen. Insgesamt lässt sich das Vorgehen also nicht einem der vier festgelegten Abläufe zuordnen, die in Abschnitt 7.1 beschrieben wurden. Vielmehr wurde an einzelnen Stellen im Forschungsprozess mit Blick auf die jeweils offenen Fragen abgewägt, welche qualitativen und quantitativen Verfahren sich an der jeweiligen Stelle anbieten.

Im Folgenden wird dieser Forschungsprozess mit dem Fokus auf die in Kapitel 6 dargelegten Forschungsfragen aufgezeigt.

F1 (*Welche herkunftsbedingten Hintergrund- und Kompetenzfaktoren haben den stärksten Zusammenhang zur Mathematikleistung in den ZP10?*) soll Aufschluss darüber geben, welchen Einfluss verschiedene Hintergrundfaktoren auf die Mathematikleistung haben. Zur Beantwortung der Frage waren quantitative Verfahren mit statistischen Auswertungsmethoden das Vorgehen der Wahl, da sich „die Forschungsfrage auf Zusammenhänge möglichst konkreter Variablen bezieht" (Hussy et al. 2013, S. 9), die standardisiert messbar sind.

Auch bei der Beantwortung von F2 (*Bei welchen Items haben viele Lernende Schwierigkeiten, bei welchen besonders die sprachlich schwachen Lernenden?*) boten sich für die Identifikation relevanter Items quantitative Auswertungsverfahren an. Diese gaben eine erste Richtung vor, von welchem ‚Typus' mathematische Prüfungsitems sein können, die sich für sprachlich schwache Lernende als besonders schwer erweisen. Zudem ermöglichte die statistische Auswertung anhand von DIF-Werten eine Fokussierung auf Items, die mittels qualitativer Verfahren genauer betrachtet werden konnten.

F3 (*Welche potenziellen sprachlichen und konzeptuellen Hürden enthalten die ZP10?*) verlangte die Realisierung eines quantitativen Vorgehens, um zu klären, inwieweit welche potenziellen Hürden in den Prüfungsitems vorhanden sind. Dabei wurde bei der Ermittlung der möglichen Hürden auf die Ergebnisse der Interviews zurückgegriffen, in denen sich auftretende Schwierigkeiten zum Teil auf bestimmte sprachliche und/oder konzeptuelle Hürden zurückführen ließen.

Auch F4 (*Bei welchen sprachlichen und konzeptuellen Hürden zeigen sich bei vielen und insbesondere bei sprachlich schwachen Lernenden Schwierigkeiten und wie sehen diese im Bearbeitungsprozess im Detail aus?*) wurde hinsichtlich einer ersten Identifikation mit einem quantitativen Vorgehen (empirische Itemanalyse) verfolgt. Damit konnten zwar Antworten zu globalen Zusammenhängen gegeben, jedoch keine Kausalitätsaussagen zur Wirkung einzelner Itemmerkmale im Bearbeitungsprozess von (sprachlich schwachen) Lernenden generiert werden. Die Erfassung solcher Bearbeitungsprozesse verlangte die Triangulation des quantitativen mit einem qualitativen Forschungsdesign, das insbesondere die Hintergründe des Zusammenhangs von Sprach- und Mathematikkompetenz rekonstruierbar machte. Die Interviews verfolgten das Ziel, Schwierigkeiten auf unterschiedlichen Ebenen untersuchen zu können, die in der Aufgliederung von F4 (vgl. Kapitel 6) benannt wurden. Die Analyse der schriftlichen Produkte (vgl. Abschnitte 16.2.3 und 16.5.1) diente hauptsächlich der Verbreiterung der Ergebnisse aus den qualitativen Analysen.

Konkreter bestand die Tiefenstudie I zunächst aus einer theoretischen Itemanalyse, in der sprachliche und mathematikdidaktische Merkmale herausgearbeitet wurden. Diese Analysen ermöglichten eine erste Auswahl von Items, die in der Interviewstudie (Tiefenstudie II) eingesetzt wurden. Die empirische Merkmalanalyse der Tiefenstudie I hatte hingegen eine andere Zielsetzung:

Durch das In-Beziehung-Setzen von Merkmalen der Items mit empirisch ermittelten absoluten und relativen Schwierigkeiten konnten einige Tendenzen aufgezeigt werden, welche Merkmale in Kombination untereinander schwierigkeitsgenerierend wirken. Das Vorgehen war dabei jedoch eher explorativ und erforderte die Triangulation mit weiteren Analysen, die Bearbeitungsprozesse sichtbar machten.

Zusammenfassend werden in der vorliegenden Arbeit quantitative und qualitativen Verfahren trianguliert, um verschiedene Facetten eines Gegenstandes zu erfassen, „die untereinander komplementär sind und sich gemeinsam zu einem vollständigeren Bild des Gegenstandes ergänzen“ (Hussy et al., S. 289).

Wie der Forschungsprozess insgesamt zeitlich ablief und aus welchen Phasen er sich zusammensetzte, wird in Kapitel 9 dargestellt. Zunächst wird der Forschungskontext dargelegt.

# 8 Forschungskontext

## 8.1 Einbindung in das übergeordnete Forschungsprojekt

Das hier behandelte Dissertationsprojekt ist eingebettet in das interdisziplinäre Forschungsprojekt „Sprachliche und konzeptuelle Herausforderungen für mehrsprachige Lernende in den Zentralen Prüfungen 10 im Unterrichtsfach Mathematik“. Beantragt wurde das Projekt als Kooperation zwischen der Mathematikdidaktik am IEEM der TU Dortmund (Susanne Prediger und Nadine Wilhelm) und dem Bereich Deutsch als Zweitsprache der Universität Duisburg-Essen, Campus Essen (Claudia Benholz und Erkan Gürsoy). Diese interdisziplinäre Kooperation ermöglichte die Integration sprachdidaktischen und mathematikdidaktischen Know-Hows in einem mehrschrittigen Forschungsprozess. Unterstützt wurde das Projekt von Andreas Büchter (jetzt Mathematikdidaktik der Universität Duisburg-Essen, Campus Essen) mit mathematikdidaktischer und psychometrischer Expertise.

Im Rahmen des zunächst einjährigen Forschungsprojekts entstand kooperativ ein Abschlussbericht (Prediger, Renk, Benholz, Gürsoy & Büchter 2013a), der die Grundlage für die Breitenauswertung der vorliegenden Arbeit darstellt. Vorgelegt wurde der Abschlussbericht im März 2013 dem Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, welches das Projekt von Januar 2012 bis Januar 2013 finanziell unterstützte. Die statistischen Analysen wurden für die Publikation Prediger et al. (2015) vertieft. Diesen Arbeitsbereich übernahm maßgeblich Andreas Büchter.

Im ersten Projektjahr wurden außerdem Interviews zur tieferen Erfassung der Bearbeitungsprozesse geführt. Die detaillierte Auswertung dieser und die

vertieften Analysen der schriftlichen Bearbeitungen wurden im Rahmen des sprachdidaktischen Dissertationsprojekts von Erkan Gürsoy und des mathematikdidaktischen Dissertationsprojekts der Verfasserin weiter geführt und gehen deutlich über den Abschlussbericht hinaus. Neben der Dissertation von Erkan Gürsoy (2013) sind Publikationen entstanden, auf die in der vorliegender Arbeit an geeigneter Stelle verwiesen wird (Gürsoy et al. 2013; Prediger, Renk, Büchter, Gürsoy & Benholz 2013b; Renk, Büchter, Prediger, Benholz & Gürsoy 2013, Gürsoy & Wilhelm 2014; Wilhelm 2014; Prediger et al. 2015).

Die Breitenstudie (vgl. Teil III) und der Großteil der Tiefenstudie I (vgl. Teil IV) wurden innerhalb der Finanzierungszeit des Ministeriums kooperativ fertig gestellt. Die Tiefenstudie II (vgl. Teil V) sowie die fachlichen und mathematikdidaktischen Analysen (vgl. Kapitel 13) stellen eigenständige Leistungen der Verfasserin dar.

Grundlage für die Breiten- und Tiefenstudien sind die Zentralen Prüfungen 10 Mathematik in Nordrhein-Westfalen, deren Rahmenvorgaben im folgenden Abschnitt 8.2 ausgeführt werden.

## 8.2 Rahmenvorgaben der Zentralen Prüfungen 10 als exemplarisch betrachtete Prüfung

Im Folgenden wird auf die Rahmenvorgaben (Relevanz, Aufbau, Ablauf, Notenvergabe) der Zentralen Prüfungen 10 Mathematik (ZP10) des Ministeriums für Schule und Weiterbildung Nordrhein-Westfalen eingegangen. Ausführungen zur inhaltlichen Konzeption sind in Abschnitt 1.3 zu finden.

Seit dem Schuljahr 2006/2007 werden der Mittlere Schulabschluss (MSA) und der Hauptschulabschluss (HSA) nach der Jahrgangsstufe 10 in Nordrhein-Westfalen in einem zentralen Abschlussverfahren vergeben. Dazu nehmen alle Schülerinnen und Schüler der Klasse 10 u.a. an Haupt-, Real- und Gesamtschulen an einem Prüfungsverfahren teil, dessen Kern eine schriftliche Prüfung mit zentral gestellten Aufgaben in den Fächern Mathematik, Deutsch und Englisch darstellt (MSW NRW 2015a). Diese Prüfungsleistung fließt zur Hälfte in die Zeugnisnote des betreffenden Faches ein. Die andere Hälfte ergibt sich aus der Vornote der jeweiligen Fachleistungen in Klasse 10. Zudem finden die Leistungen der übrigen Fächer ohne zentrale Prüfungen Berücksichtigung bei der Abschlussvergabe. Für die Fächer Mathematik, Deutsch und Englisch spielen die ZP10 somit eine enorme Rolle bei der Erreichung des angestrebten Schulabschlusses. Aufgrund der bereits benannten Ausrichtung dieser Arbeit beziehen sich die folgenden Ausführungen ausschließlich auf das Fach Mathematik.

Die ZP10 beinhalten zwei Prüfungsteile. In Teil 1 (Item 1a–1e) werden sogenannte „Basiskompetenzen“ (MSW NRW 2015a) in einzelnen, nicht aufeinander bezogenen Teilaufgaben thematisiert. Diese beziehen sich auf Kompetenzen, die im Laufe der Jahrgangsstufen 5 bis 10 aufgebaut werden sollen. Für

das Fach Mathematik sind 30 Minuten zur Bearbeitung des ersten Prüfungsteils vorgesehen. Das Item 1a Münzturm ist als eine von fünf dieser Teilaufgaben des Prüfungsjahrgangs 2012 in Abbildung 8.1 abgedruckt. Die Abbildung dient hier exemplarisch zur Illustration. Aufgrund der besonderen Anforderungen dieses Items für sprachlich schwache Lernende findet sich in Kapitel 13 eine detaillierte Analyse.

a) Schätze, wie viele Kilometer hoch ein Turm aus 2,4 Milliarden 1-Cent-Münzen ungefähr wäre. Beschreibe, wie du vorgegangen bist.

Abbildung 8.1 Item 1a Münzturm ZP10 2012

Im zweiten Prüfungsteil (Item 2–4) werden Aufgaben in drei komplexeren außermathematischen Kontexten mit jeweils mehreren Teilaufgaben gestellt, zu deren Bearbeitung sowohl prozess- als auch inhaltsbezogenes Arbeiten erforderlich ist (zur vollständigen Fassung der Prüfung des Jahrgangs 2012 vgl. MSW NRW 2012a, zum Korrekturleitfaden MSW NRW 2012b). Grundlage für diesen Prüfungsteil sind Kompetenzen, die im Doppeljahrgang 9/10 erworben wurden. Zur Bearbeitung des zweiten Prüfungsteils stehen den Schülerinnen und Schülern 90 Minuten zur Verfügung.

Auch aus dem zweiten Prüfungsteil sollen hier einige Items die Gestalt der Prüfungsaufgaben deutlich machen. Alle dargestellten Items zeigen sich in der Breiten- und/oder in den Tiefenstudien als auffällig und werden in Kapitel 13 detailliert analysiert.

Der Kraftstoffverbrauch wird für Fahrzeuge durch den durchschnittlichen Verbrauch in Litern (ℓ) auf einer Strecke von 100 Kilometern angegeben. Der Kraftstoffverbrauch eines Autos hängt vor allem von der gefahrenen Geschwindigkeit ab.

a) Das Diagramm zeigt den Kraftstoffverbrauch für ein Auto, das im höchsten Gang gefahren wird. Daher beginnt der Graph bei 70 km/h.

   (1) Wie schnell fährt das Auto durchschnittlich, wenn es 11 ℓ auf 100 km verbraucht?

   (2) Um wie viel Prozent liegt der Verbrauch bei 180 km/h über dem Verbrauch bei 100 km/h? Notiere deine Rechnung.

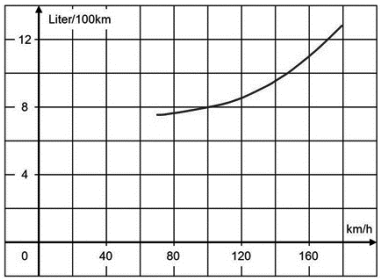



Abbildung 8.2 Item 2a1 Verbrauch-ablesen und Item 2a2 Verbrauch-Prozent ZP10 2012 (hier zusammen dargestellt aufgrund des notwendigen Zugriffs auf das Diagramm beider Items)

b) Familie Wacker fährt mit einem vollgetankten Auto in den Urlaub. Ihr Fahrzeug hat einen Bordcomputer, der während der Fahrt u. a. Informationen über gefahrene Kilometer und Kraftstoffverbrauch berechnet und anzeigt (siehe Tabelle).

| gefahrene Kilometer | Anzeige | | |
|---|---|---|---|
| | Verbrauch in Litern | durchschnittlicher Verbrauch (in ℓ/100 km) | verbleibende Reichweite (in km) |
| 180 | 14,6 | | 485 |

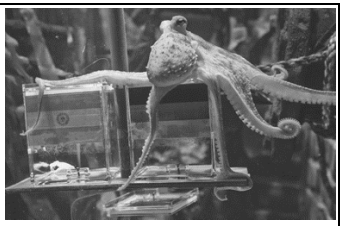
(1) Nach 180 km und einem Verbrauch von 14,6 ℓ Kraftstoff macht Familie Wacker eine erste Pause. Zeige, dass das Auto bis zur ersten Pause einen durchschnittlichen Verbrauch von 8,1 ℓ/100 km hatte.

Abbildung 8.3 Item 2b1 Verbrauchsberechnung ZP10 2012

c) Für das Auto von Familie Wacker lässt sich der durchschnittliche Kraftstoffverbrauch (in ℓ/100 km) in Abhängigkeit von der Geschwindigkeit (in km/h) näherungsweise mit der folgenden Gleichung berechnen:
$f(x) = 0{,}0005 \cdot (x - 40)^2 + 4{,}5462$.

(1) Wie hoch ist der durchschnittliche Verbrauch bei einer Geschwindigkeit von 150 km/h? Notiere deine Rechnung.

Abbildung 8.4 Item 2c1 Verbrauch-y ZP10 2012



Bei der Fußball-WM 2010 wurde der Krake Paul international berühmt. Vor jedem Fußballspiel wurden zwei Futterboxen in sein Aquarium gesenkt. Die Boxen waren mit der jeweiligen Flagge der beiden Länder beklebt, deren Mannschaften gegeneinander spielten. Paul suchte sich einen der beiden Futtertöpfe aus. Seine Wahl wurde dann von den Medien als „Vorhersage“ des Gewinners des Fußballspiels gedeutet. Da Paul alle Spiele der deutschen Nationalmannschaft richtig voraussagte, wurde er ein richtiger Medienstar.

Gehe davon aus, dass Pauls „Vorhersagen“ zufällig geschehen sind. Mathematisch betrachtet, handelt es sich bei den „Vorhersagen“ also um einen Zufallsversuch mit zwei gleich wahrscheinlichen Ergebnissen.

a) Erkläre, wie man diesen Zufallsversuch mithilfe eines Würfels simulieren kann.

Abbildung 8.5 Item 3a Simulation ZP10 2012

b) Zeichne ein Baumdiagramm, das die Wahrscheinlichkeiten für zwei Vorhersagen angibt.

c) Zeige, dass die Wahrscheinlichkeit, dass Paul zwei Spiele hintereinander richtig tippt, $\frac{1}{4}$ beträgt.

Abbildung 8.6 Item 3b Baumdiagramm und 3c Wahrscheinlichkeit-Baum ZP10 2012 (hier zusammen dargestellt aufgrund der engen Verzahnung, vgl. Kapitel 13)

Neben der Gesamtbearbeitungszeit der beiden Prüfungsteile von 120 Minuten haben Lernende zusätzlich zehn Minuten Zeit zur Orientierung. Die Fachlehrkraft entscheidet, für welchen Prüfungsteil diese Zeit eingesetzt wird, wobei auch eine Aufteilung in jeweils fünf Minuten möglich ist.

Die Bewertung der Bearbeitungen erfolgt nach einem dreischrittigen Schema, das folgendermaßen zusammengefasst werden kann (zur näheren Ausführung und zur Diskussion der Rohpunktvergabe vgl. Büchter & Pallack 2012, S. 64ff):

Bearbeitung → Auswertungskriterien → Rohpunkte → Notentabelle

Durch jeweils zwei Fachlehrkräfte werden die Lernendenleistungen anhand halb-abstrakt formulierter, möglichst standardisierter Auswertungskriterien erfasst (MSW NRW 2012b). Anschließend werden ganze Rohpunkte für die einzelnen Kriterien je nach Grad der Erfüllung bis zu einer vorgegebenen Maximalpunktzahl vergeben (ausführlicher vgl. Büchter & Pallack 2012). In Abstimmung mit der Lehrkraft, welche die Zweitkorrektur durchgeführt hat, setzt die Fachlehrkraft die Note fest. Falls keine Einigung stattfindet, zieht die Schulleitung eine dritte Lehrkraft hinzu und die Note wird durch Mehrheitsbeschluss festgelegt.

Im Fach Mathematik werden zudem der Umgang mit Maßeinheiten und die Nachvollziehbarkeit der Darstellung zusammenfassend bewertet und gesondert im Bewertungsbogen ausgewiesen. Eine Übersicht der Punkteverteilung in den ZP10 des Prüfungsjahrgangs 2012 aus der Handreichung für Lehrkräfte (MSW NRW 2012b) findet sich in der Tabelle 8.1.

Tabelle 8.1 Übersicht über die Punkteverteilung

| Aufgabe bzw. Leistung | | Punkte |
|---|---|---|
| Prüfungsteil 1 | Aufgabe 1 | 19 |
| Prüfungsteil 2 | Aufgabe 2 | 17 |
| | Aufgabe 3 | 22 |
| | Aufgabe 4 | 18 |
| Umgang mit Maßeinheiten | | 03 |
| Darstellungsleistung | | 06 |
| Gesamtpunktzahl | | 85 |

Der letzte Schritt des oben dargestellten dreistufigen Schemas stellt die Ermittlung der Note gemäß der erreichten Gesamtpunktzahl dar. Den Notenschlüssel für den Durchgang 2012 zeigt Tabelle 8.2.

Tabelle 8.2 Notentabelle

| Punkte | Note |
|---|---|
| 74–85 | sehr gut |
| 62–73 | gut |
| 50–61 | befriedigend |
| 38–49 | ausreichend |
| 15–37 | mangelhaft |
| 00–14 | ungenügend |

Die Bedeutung der Punkte in den ZP10 für die Analysen der vorliegenden Arbeit wird in Abschnitt 10.1.1 dargelegt. Zunächst wird ein Überblick über den zeitlichen Auflauf und die Phasen des Dissertationsprojekts gegeben.

# 9 Überblick über den zeitlichen Ablauf und die Phasen des Dissertationsprojekts

Den Breiten- und Tiefenstudien liegen mehrere Phasen der Datenerhebung und -auswertung zugrunde, die im Folgenden überblicksartig vorgestellt werden. Innerhalb der Phasen war stets das Untersuchungsinteresse handlungsleitend für die jeweils spezifische Methodik. Die Phasen werden hier zwar einzeln aufgelistet (vgl. Tabelle 9.1) und dargestellt, standen im Forschungsprozess aber in einem engen Zusammenhang. So ergaben sich durch einzelne Phasen entweder neue Fragen oder es wurde erkannt, dass zur Beantwortung von Fra-

gen verschiedene Zugriffe notwendig waren (für Details dieses Prozesses vgl. Abschnitt 7.2).

Tabelle 9.1 zeigt die Phasen des Dissertationsprojektes im Überblick, wobei die Phasen V bis VII der Breitenstudie und die Phasen I bis IV sowie VIII bis XIV den Tiefenstudien zuzuordnen sind.

Tabelle 9.1 Phasen des Dissertationsprojektes und Zeiträume der Umsetzung

| **Phasen** | | **Zeitraum (Quartal und Jahr)** |
|---|---|---|
| I | Theoretische Itemanalyse von 54 Items | 1 (2012) – 2 (2012) |
| II | Vorbereitung der ersten Interviewserie | 1 (2012) |
| III | Durchführung der ersten Interviewserie mit 29 Lernenden in 23 Interviews | 1 (2012) – 3 (2012) |
| IV | Grobe Auswertung der ersten Interviewserie | 2 (2012) – 4 (2012) |
| V | Vorbereitung der Breitenerhebung | 1 (2012) |
| VI | Durchführung der Breitenerhebung mit 1495 Lernenden | 2 (2012) – 3 (2012) |
| VII | Auswertung der Breitenerhebung | 3 (2012) – 4 (2012) |
| VIII | Empirische Itemanalyse von 27 Items | 4 (2012) |
| IX | Vorbereitung der Analyse der schriftlichen Produkte | 3 (2012) – 4 (2012) |
| X | Analyse von 195 schriftlichen Produkten | 1 (2013) – 2 (2013) |
| XI | Vorbereitung der zweiten Interviewserie | 1 (2013) |
| XII | Durchführung der zweiten Interviewserie mit 13 Lernenden in 13 Interviews | 2 (2013) |
| XIII | Auswertung der zweiten Interviewserie | 3 (2013) – 1 (2014) |
| XIV | Vertiefte Auswertung der ersten Interviewserie | 3 (2013) – 4 (2013) |
| XV | Schreibphase (unterbrochen durch Elternzeit) | 4 (2013) – 4 (2015) |

Wie Tabelle 9.1 entnommen werden kann, überlagern sich die einzelnen Phasen zum Teil zeitlich. Zur Schaffung von Transparenz über den inhaltlichen Forschungsverlauf werden im Folgenden die Phasen dargestellt. Für Details wird auf die entsprechenden Abschnitte verwiesen.

*Phase I: Theoretische Itemanalyse (vgl. Abschnitt 14.1)*

In einem ersten Zugriff wurden die 27 Prüfungsitems der ZP10 aus dem Jahrgang 2011 und die 27 Items aus dem Jahrgang 2012 hinsichtlich ihrer mathematikdidaktischen und sprachlichen Charakteristika und potenziellen Hürden analysiert. Insgesamt sollten so mehrere Ziele erreicht werden: Zu Beginn des Projekts ging es insbesondere darum, allgemeine Merkmale des Prüfungsformats ZP10 erfassen zu können. Theoriegeleitet wurden dann a priori Hypothesen aufgestellt, welche Itemmerkmale (im Zusammenspiel) potenziell

schwierigkeitsgenerierend für alle und insbesondere für sprachlich schwache Lernende wirken können.

In den weiteren Forschungsphasen dienten die theoretisch spezifizierten potenziell schwierigkeitsgenerierenden Merkmale als Grundlage für die empirische Itemanalyse, um die empirische Bedeutung der Merkmale für die tatsächliche Itemschwierigkeit zu erfassen. Zudem sensibilisierte die Aufstellung der Merkmale für die Interviews in Bezug auf die Auswahl der Items (*Phase II* und weniger stark aufgrund weiterer Zugriffe und vorangegangener Analysen *Phase XI*) und der Durchführung (*Phase III* und *XII*), d.h. der Fokussetzung der Fragestellungen. Auch für die Analyse der schriftlichen Produkte (*Phase X*) stellte die theoretische Itemanalyse die Basis dar.

*Phase II: Vorbereitung der ersten Interviewserie (vgl. Abschnitt 15.2.1)*

Auf der Grundlage der theoretischen Itemanalyse wurden Items der Prüfungsjahrgänge 2011 (hier betrachtet 4c1 WM-Eintrittspreis und 4c2 WM-Zuschauerschnitt) und 2012 (1b Bakterienverdopplung, 2a1 Verbrauch-ablesen, 2a2 Verbrauch-Prozent, 2b1 Verbrauchsberechnung und 2c1 Verbrauch-y) ausgewählt, die in halbstandardisierten leitfadengestützten Interviews eingesetzt werden sollten.

Rekrutiert wurden die Interviewteilnehmerinnen und -teilnehmer im Förderunterricht für Kinder und Jugendliche mit Migrationshintergrund an der Universität Duisburg-Essen, Campus Essen. Da sich der Förderunterricht nicht auf einzelne Schulformen beschränkt, waren neben größtenteils Gesamtschullernenden ebenfalls Haupt-, Real- und Berufskolleglernende Interviewteilnehmerinnen und -teilnehmer, die größtenteils die 10. Klasse besuchten und 15–17 Jahre alt waren.

*Phase III: Durchführung der ersten Interviewserie (vgl. Abschnitt 15.2.1)*

In halbstandardisierten klinischen Interviews, die größtenteils jeweils 45 bis 60 Minuten dauerten, wurden Bearbeitungsprozesse von Lernenden erhoben. Bis zum Prüfungsdatum des Jahres 2012 wurden in zehn Einzelinterviews Items des Jahrgangs 2011 eingesetzt, nach dem offiziellen Prüfungstermin die des Jahrgangs 2012 in weiteren sieben Einzelinterviews und sechs Partnerinterviews.

*Phase IV: Grobe Auswertung der ersten Interviewserie (vgl. Abschnitt 15.3.1)*

In einer ersten Durchsicht der rund 1050 Minuten Videomaterial wurden in Abgleich mit der theoretischen Itemanalyse tatsächliche Hürden in den Bearbeitungsprozessen identifiziert und einzelne Interviewepisoden transkribiert. Zur ersten Spezifizierung wurden die beobachteten Schwierigkeiten groben Ebenen (wie z.B. Wort-, Satz- und Textebene im Leseverstehensprozess der Items) zugeordnet. Aus den Ergebnissen wurden die durch die theoretische

Itemanalyse aufgestellten Hypothesen angepasst und erweitert, da neue Schwierigkeiten sichtbar wurden.

*Phase V: Fertigstellung der Erhebungsinstrumente für die Breitenerhebung (vgl. Abschnitt 10.1)*

Zur Vorbereitung der Breitenerhebung wurden die Testinstrumente und Fragebögen zur Erfassung der Sprachkompetenz (C-Test) und der herkunftsbedingten Hintergrundfaktoren (u.a. sozioökonomischer Status und Sprachhintergrund, erhoben durch einen Sprachbiographiebogen) zusammengestellt. Die Mathematik- und Lesekompetenz wurde anhand der Bearbeitungen der zentral gestellten landeseinheitlichen Aufgaben gemessen und bedurfte keiner Vorbereitung, da die Prüfungen durch die Schulen durchgeführt wurden (vgl. Phase VI).

*Phase VI: Durchführung der Breitenerhebung (vgl. Abschnitt 10.1)*

Zeitlich parallel zur Durchführung und Auswertung der ersten Interviewserie fand die Breitenerhebung an 19 Gesamtschulen mit insgesamt 67 Mathematik-Erweiterungskursen (ein für das Ruhrgebiet repräsentatives Sample) statt, die Stichprobe umfasste n = 1495 Testbögen.

Die Zentralen Prüfungen Deutsch und Mathematik wurden unter normalen Feldbedingungen einer abschlussrelevanten Prüfung am zentral für Nordrhein-Westfalen festgelegten Termin von allen Lernenden geschrieben und durch die Fachlehrkräfte entlang des standardisierten Bewertungsbogens korrigiert. Die Prüfungsergebnisse von 1495 Lernenden an 19 Gesamtschulen wurden im Anschluss an das Projektteam weitergeleitet. Die Fragebögen zur Erfassung der herkunftsbedingten Hintergrundfaktoren von 1493 Lernenden und der Sprachkompetenz von 698 Lernenden wurden kurz vor oder kurz nach den ZP10 in den Projektschulen durch die Verfasserin oder geschulte Mitarbeiterinnen und Mitarbeiter erhoben.

Ergebnisse zur Sprachkompetenz, gemessen durch den C-Test, lagen von 698 Lernenden vor, da mit Blick auf den testbedingten Unterrichtsausfall kurz vor den Zentralen Prüfungen nur ein Teil der Schulen an dieser zusätzlichen Erhebung teilgenommen hat. Die Mathematikleistung dieser Teilstichprobe ($M = 11{,}1$; $SD = 4{,}7$, vgl. Tabelle 11.8 in Abschnitt 11.1.1) unterscheidet sich aber nicht signifikant von der Mathematikleistung der Gesamtstichprobe. Hinsichtlich der Lesekompetenz wurde der Erweiterungskurs Deutsch als Stichprobe mit 1066 Lernenden ausgewählt. Grund dafür war die weitaus höhere Anzahl von Lernenden, die in die Auswertung einfließen konnten (zur Begründung vgl. auch Abschnitt 10.1.3).

Die Ergebnisse sämtlicher Erhebungen wurden von Mitarbeiterinnen und Mitarbeiter der TU Dortmund eingegeben und anschließend von der Verfasserin zur statistischen Auswertung aufbereitet.

*Phase VII: Auswertung der Breitenerhebung (vgl. Abschnitt 10.3)*

Mittels verschiedener statistischer Verfahren wurden Hintergrundfaktoren und Mathematikleistung bzw. Item-Schwierigkeit miteinander in Beziehung gesetzt. Einerseits wurden deskriptive Auswertungen und Darstellungen auf der Ebene von absoluten und relativen Häufigkeiten verwendet und anderseits im Anschluss an eine Rasch-Skalierung der Prüfungsdaten Varianz-, Regressions- und DIF-Analysen durchgeführt.

*Phase VIII: Empirische Itemanalyse (vgl. Abschnitt 14.1)*

Die in der theoretischen Itemanalyse (vgl. Phase I) ermittelten potenziell schwierigkeitsgenerierenden Merkmale wurden durch Erkenntnisse der Interviews überarbeitet. Dabei wurden insbesondere Merkmale ausgeschärft, die in Phase I in allgemeinerer Form formuliert wurden. Anschließend wurden die angepassten Merkmale mit den empirisch ermittelten absoluten und relativen Schwierigkeiten in Beziehung gesetzt. Aufgrund der geringen Itemanzahl (n = 27 des Jahrgangs 2012) wurde dabei eher explorativ vorgegangen. Es konnten Tendenzen für schwierigkeitsgenerierende Merkmale für alle Lernende und insbesondere sprachlich Schwache identifiziert werden.

*Phase IX: Vorbereitung der Analyse der schriftlichen Produkte (vgl. Abschnitte 15.2.2 und 15.3.2)*

Für Items, die sich in den vorangegangenen Analysen als besonders interessant und auffällig in Bezug auf die DIF-Werte und typische schwierigkeitsgenerierende Item-Merkmale herausstellten, wurden Analyse-Kategorien (i.S. von Bearbeitungsschritten) für die Auswertung von 195 schriftlichen Produkten (Klausurverschriftlichungen) der Schülerinnen und Schüler aufgestellt. Das jeweils aufgabenspezifisch entwickelte Kategoriensystem berücksichtigt sowohl sprachlich rezeptive als auch sprachlich produktive sowie fachliche Elemente.

*Phase X: Analyse der schriftlichen Produkte (vgl. Abschnitt 15.3.2)*

195 schriftliche Bearbeitungen von Schülerinnen und Schülern wurden entlang der aufgabenspezifischen Kategoriensysteme analysiert. Mittels statistischer Verfahren wurden Lösungshäufigkeiten und Bearbeitungsprofile für Gruppen mit unterschiedlicher Sprachkompetenz verglichen.

*Phase XI: Vorbereitung der zweiten Interviewserie (vgl. Abschnitt 15.2.1)*

Ausgangspunkt für eine zweite Interviewserie waren zum einen die Ergebnisse der Breitenstudie, insbesondere der DIF-Analysen: Dabei stellten sich einige Items als relativ schwieriger für sprachlich Schwache dar, die durch die theoretische Itemanalyse nicht in den Fokus gerückt waren. Zum anderen sollten ge-

zieltere Einblicke in Bearbeitungsprozesse von Items geliefert werden, für deren Analyse die schriftlichen Produkte nicht ausreichend aufschlussreich waren.

Da zum Zeitpunkt der zweiten Interviewserie (zweites Quartal 2013) die Lernenden der ersten Interviewserie und des Gesamtsamples der Breitenstudie nicht mehr erreichbar waren, wurde eine Jahrgangsstufe 10 einer weiteren Gesamtschule ausgewählt.

*Phase XII: Durchführung der zweiten Interviewserie (vgl. Abschnitt 15.2.1)*

Die Durchführung der zweiten Interviewserie fand ähnlich wie die erste Interviewserie statt. In Einzelinterviews mit 13 Lernenden wurden im Kontrast zur ersten Interviewserie andere Items mit dem Fokus auf (vermutete) konzeptuelle Hürden eingesetzt. Die Interviews wurden durch die Verfasserin alleine durchgeführt. Die Sensibilität für sprachliche Merkmale der eingesetzten Items wurde durch die Einbeziehung der Ergebnisse der theoretischen Itemanalyse gewährleistet.

*Phase XIII: Auswertung der zweiten Interviewserie (vgl. Abschnitt 15.3.1)*

Ähnlich der Auswertung der ersten Interviewserie wurden zunächst auftretende Schwierigkeiten in den rund 340 Minuten Videomaterial grob erfasst und interessante Stellen transkribiert. Zur genaueren und methodisch abgesicherten Verortung von tatsächlichen Schwierigkeiten und sichtbaren (tragfähigen) Vorstellungen wurden aufgabenbezogene Konstrukte aus fachlicher und individueller Perspektive der Lernenden spezifiziert. Die Interviews der zweiten Interviewserie wurden anhand der Konstrukte qualitativ in einem kategorienentwickelnden Verfahren analysiert.

*Phase XIV: Vertiefte Auswertung der ersten Interviewserie (vgl. Abschnitt 15.3.1)*

Auf der Grundlage gewonnener Ergebnisse aus vorangegangenen Forschungsschritten wurden itemspezifische Analysewerkzeuge auch für die Daten der ersten Interviewserie entwickelt, um ausgewählte Items und Transkripte vertieft analysieren zu können. Interviewepisoden, in denen Schwierigkeiten des Aufgabenverstehens auf Wort-, Satz- und Textebene hervor traten, wurden hauptsächlich mit Hilfe des Leseprozessmodells (vgl. Abschnitt 4.2.1) und mit Reussers Prozessmodell (1992, vgl. Abschnitt 5.1) analysiert. Andere beobachtbare Schwierigkeiten (beim Bilden des Situationsmodells und auf konzeptueller Ebene) erforderten die Entwicklung von neuen Analysemodellen.

Die innerhalb der Phasen konkret eingesetzten Methoden werden in den folgenden empirischen Kapiteln der Breiten- und Tiefenstudien erläutert. Das folgende Kapitel geht zunächst auf die Breitenstudie ein.

http://www.springer.com/978-3-658-13735-9

Zusammenhänge zwischen Sprachkompetenz und Bearbeitung mathematischer Textaufgaben
Quantitative und qualitative Analysen sprachlicher und konzeptueller Hürden

Wilhelm, N.

2016, XVI, 320 S. 25 Abb., Softcover
ISBN: 978-3-658-13735-9