



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

Fehlerkategorisierung in Ausarbeitungen der
standardisierten schriftlichen Reifeprüfung Mathematik

Eine Analyse des Haupttermins 2016 an AHS

verfasst von / submitted by

Johannes Klien

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 190 333 406

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium UF Deutsch UF Mathematik

Betreut von / Supervisor:

Ao. Univ.-Prof. Mag. Dr. Stefan Götz

Danksagung

An dieser Stelle möchte ich mich bei all jenen Personen bedanken, die mich während meines Studiums begleitet, unterstützt und motiviert haben. Insbesondere bedanke ich mich bei

... Herrn Ao. Univ.-Prof. Mag. Dr. Stefan Götz, durch dessen Betreuung diese Arbeit erst ermöglicht wurde und der mir mit seinem informativen Feedback sehr geholfen hat. Vielen Dank für die ausgezeichnete Zusammenarbeit!

... Frau Mag. Dr. Eva Sattlberger, Herrn Mag. David Stadler vom Bundesministerium für Bildung, Wissenschaft und Forschung, die als Ansprechpersonen stets ein offenes Ohr hatten und immer hilfreich zur Stelle waren sowie dem Bundesministerium für Bildung, Wissenschaft und Forschung für die Genehmigung mit den Daten arbeiten zu dürfen.

... meinen Freunden Julia, Jess, Josef, Katrin, Keuz, Lina, Tina und Vicky, mit denen ich während des Studiums viele angenehme Stunden verbringen durfte.

... meinen Freunden und Mitbewohnern Ivo, Mät und Chris – gemeinsam haben wir unsere Studienzeit sehr genossen.

... meinen Eltern, meinen Schwestern, meinen Omas, meiner Familie, auf deren Unterstützung ich immer zählen konnte.

Vielen Dank!

Inhaltsverzeichnis

1	Einleitung.....	1
2	Grundlagen der Datenanalyse.....	3
2.1	Daten.....	3
2.1.1	Stichprobe	3
2.1.2	Auswahl.....	5
2.2	Prozesse empirischer Forschung.....	13
2.2.1	Erhebung	15
2.2.2	Aufbereitung	17
2.2.3	Auswertung	19
3	Methodik	27
3.1	Induktive Kategorienbildung.....	28
3.2	Häufigkeitsanalyse	35
3.3	Exemplarische Auswertung	37
4	Auswertung	46
4.1	Aufgabe 1	48
4.2	Aufgabe 2	53
4.3	Aufgabe 3	56
4.4	Aufgabe 4	59
4.5	Aufgabe 5	61
4.6	Aufgabe 6	62
4.7	Aufgabe 19.....	66
4.8	Aufgabe 20.....	69
4.9	Aufgabe 21.....	74
4.10	Aufgabe 22	77
4.11	Aufgabe 23	79

4.12	Aufgabe 24	83
5	Resümee und Ausblick.....	87
	Literaturverzeichnis.....	89
	Abstract	93

1 Einleitung

Nach der Einführung der standardisierten schriftlichen Reifeprüfung (ssRP) in Mathematik im Schuljahr 2014/15 wurde im Schuljahr 2015/16 der nächste Schritt für die Zentralmatura Mathematik eingeleitet. In diesem Jahr wurde zum ersten Mal eine Stichprobe der Performanzen der ssRP erhoben. Dies ermöglicht neue Erkenntnisse und vor allem die wissenschaftliche Arbeit mit konkreten Ausarbeitungen von Schülerinnen und Schülern. Die vorliegende Arbeit bearbeitet diese erstmals zum Haupttermin 2016 vom Bundesministerium für Bildung (BMB) erhobenen Daten.

Das Hauptaugenmerk wird dabei auf die Ausarbeitungen der Schülerinnen und Schüler gelegt. Die Arbeiten werden Aufgabe für Aufgabe durchgegangen und die verschiedenen Lösungsansätze (egal ob richtig oder falsch) klassifiziert. Die größere Bandbreite an verschiedenen Lösungen findet sich erwartungsgemäß unter den falsch beantworteten Aufgaben. Aus diesem Grund werden Kategorien innerhalb von fehlerhaft bearbeiteten Aufgaben gebildet. So soll am Ende zu jeder untersuchten Aufgabe ein System an Kategorien stehen, das es ermöglicht, mit diesem Detailwissen über die Lösungsvielfalt von Aufgaben weiterführend zu arbeiten. Das Hauptziel ist es also die vorhandenen Rohdaten nutzbar zu machen.

Dazu wird in *Kapitel 2* zunächst die Datenlage vorgestellt. Die Daten werden in weiterer Folge detailliert beschrieben, erste Kennzahlen werden genannt. Der Prozess der Auswahl wird daraufhin beschrieben und so nachvollziehbar gemacht. Im zweiten Abschnitt dieses Kapitels wird ein Überblick über den empirischen Arbeitsablauf, ausgehend von Erhebung, Aufbereitung und Auswertung, gegeben.

Kapitel 3 stellt die verwendeten Methoden vor und bereitet so die Auswertung vor. Die nach der induktiven Kategorienbildung erstellten Kategorien werden mithilfe einer Häufigkeitsanalyse weiter ausgewertet. Es werden somit qualitative und quantitative Ansätze verwendet. Eine ausführliche Beschreibung des Arbeitsprozesses anhand eines konkreten Beispiels findet sich unter *Abschnitt 3.3 Exemplarische Auswertung* wieder.

Im Kernstück der Arbeit, *Kapitel 4 Auswertung*, werden schließlich die zwölf ausgewählten Aufgaben Schritt für Schritt beschrieben und nach den in *Kapitel 3* vorgestellten Methoden bearbeitet. Dabei entsteht zu jeder einzelnen Aufgabe ein System an Kategorien,

das zunächst wiederum, auch mithilfe anonymisierter Performanzen, beschrieben wird und im Anschluss daran mittels einer Häufigkeitsanalyse quantitativ ausgewertet wird. Mit den jeweiligen Kategoriensystemen und den Häufigkeitstabellen sowie Statistiken gilt die deskriptive Datenanalyse als abgeschlossen.

2 Grundlagen der Datenanalyse

Am Beginn jeder Arbeit steht ein Überblick über das zu bearbeitende Feld und der Klärung des Ziels der Forschung (RIESENHUBER 2009: 4). Im vorliegenden Fall geht es vor allem darum die vorhandenen Daten nutzbar zu machen, das heißt eine Auswahl zu treffen und diese wiederum detaillierter zu bearbeiten. Die Kategorien, die dabei entstehen, werden durch systematisches Vorgehen (näheres zur Methode in *Kapitel 3*) nach und nach erarbeitet, so dass am Ende gesicherte wissenschaftliche Erkenntnisse stehen.

Um einen ersten Überblick über den Untersuchungsgegenstand zu erhalten, werden zuerst die Daten, die vom Bundesministerium für Bildung 2017 zur Verfügung gestellt wurden, genauer vorgestellt. Danach wird mithilfe von HUG & POSCHESCHNIK (2015) ein grundsätzlicher Überblick geschaffen, bevor die in weiterer Folge verwendeten Methoden näher betrachtet werden.

2.1 Daten

Als Untersuchungsgegenstand wurde die standardisierte schriftliche Reifeprüfung Mathematik aus 2016 gewählt. Aus den möglichen drei Terminen, Haupttermin 2015/16 sowie den beiden Nebenterminen, wurde der Haupttermin gewählt. Diese Auswahl wurde aus folgendem Grund getroffen: Zum Zeitpunkt des Beginns der Untersuchungen handelte es sich bei den Daten des Haupttermins 2016 um den aktuellsten vollständigen Datensatz.

2.1.1 Stichprobe

Das Bundesministerium für Bildung erhebt seit 2015 Daten zur standardisierten schriftlichen Reifeprüfung Mathematik an den allgemein bildenden höheren Schulen (AHS). Bei der ersten Durchführung der ssRP Mathematik im Jahr 2015 wurde die Punkteverteilung an den Schulstandorten erhoben. Im folgenden Jahr wurden beim Haupttermin erstmals die Performanzen einer Prüfungsgruppe digitalisiert. Dazu wurde nach Abschluss des Termins eine Stichprobe erhoben. Diese Stichprobe enthält die Performanzen der Schülerinnen und Schüler, also die konkreten Arbeiten (Hefte) zur Matura. Insgesamt werden dabei 1000 Performanzen erhoben, pro Standort aber nur die Performanzen von sechs Kandidatinnen oder Kandidaten. Dies ergibt eine Auswahl von 166 Schulstandorten ($1000/6 = 166,67$). In weiterer Folge werden zunächst die Schulstandorte pro Bundesland ermittelt, um zu berechnen wie viele Schulstandorte anteilmäßig pro Bundesland ausgewählt werden

müssen. Beispielsweise sollen burgenländische Schulstandorte drei Prozent von 166 Standorten ausmachen, deshalb wurden fünf Standorte aus dem Burgenland gewählt.

Innerhalb jedes Bundeslandes wird die Häufigkeit der einzelnen Schulformen berücksichtigt. Kommt eine Schulform in einem Bundesland häufiger vor, so schlägt sich das auch in der Stichprobe nieder. Nachdem ermittelt wurde, wie sich die Verteilung der verschiedenen Schulformen innerhalb eines Bundeslandes gestaltet, werden die Schulstandorte, wiederum anteilig, zufällig ausgewählt. Befinden sich in einem hypothetischen Bundesland 100 Schulen, von denen 22 der Schulform A, 33 der Schulform B, 40 der Schulform C und 5 der Schulform D zugerechnet werden, dann werden auch 22 Prozent der Stichprobe aus diesem Bundesland aus der Schulform A gezogen.

Wenn die einzelnen, zufällig ausgewählten, Schulstandorte ermittelt wurden, wird dort eine Klasse zufällig gewählt, aus der wiederum zufällig maximal sechs Kandidatinnen und Kandidaten ausgewählt werden. Bei dieser Stichprobe handelt es sich also um eine mehrstufige, gewichtete Zufallsstichprobe¹. Die Datengrundlage, auf der diese Stichprobenziehung basiert, stellen die sogenannten Drucklisten dar. Dies sind die Anmeldungen der Schülerinnen und Schüler zur ssRP.

Nach Aufzeichnungen des BMB wurden in Mathematik mit der oben beschriebenen Auswahlmethode 1002 Kandidatinnen und Kandidaten zufällig ermittelt. Diese zufällig ermittelten Kandidatinnen und Kandidaten wurden also repräsentativ für die Bundesländer und die Schulformen innerhalb der Bundesländer gezogen. 878 Hefte erreichten das BMB, die Rücklaufquote beträgt also etwa 87,6 Prozent. Diese vom BMB anonymisierten und digitalisierten Performanzen bilden die Grundlage dieser Arbeit.

Der Prozess der Stichprobenziehung wird in Abbildung 1 nochmals zusammengefasst:

¹ Bei einer mehrstufigen Zufallsstichprobe werden die Untersuchungsobjekte auf mehreren Ebenen zufällig ausgewählt (HÄDER & HÄDER 2014: 285). Laut KAYA & HIMME (2009: 85) kommt „dieses Verfahren [typischerweise] bei hierarchisch gegliederten Grundgesamtheiten“ zum Einsatz. Hier heißt das, dass auf der Ebene der Schulformen pro Bundesland zunächst zufällig Schulstandorte und im Anschluss daran, auf der Ebene der Schulen, zufällig Klassen ausgewählt werden.

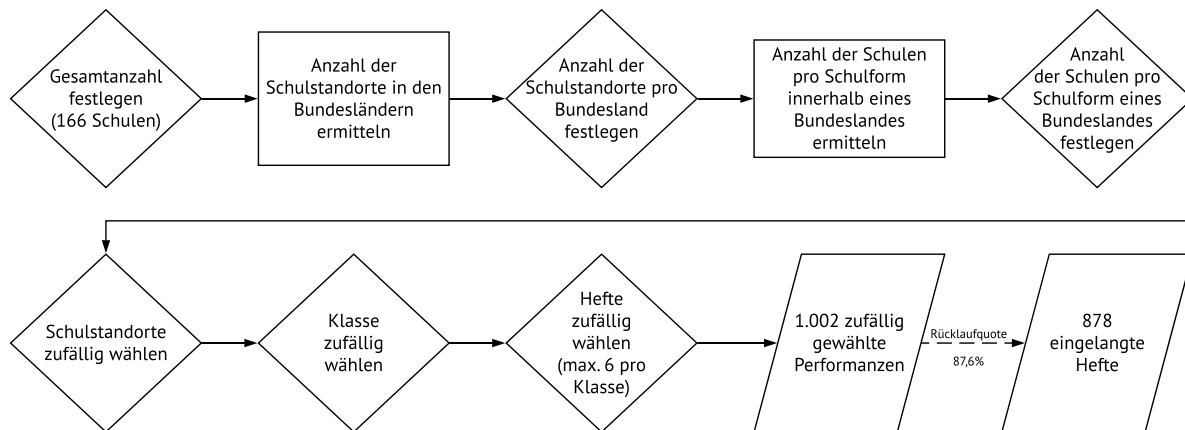


Abbildung 1: Stichprobenziehung

2.1.2 Auswahl

Die standardisierte schriftliche Reifeprüfung Mathematik besteht aus zwei Teilen. Der erste Teil, der die sogenannten Typ-1-Aufgaben umfasst, beinhaltet 18 bis 25 Aufgaben. De facto bestand jeder erste Teil aller bisherigen ssRP Mathematik (2013/14-2015/16, inklusive aller Nebentermine) aus 24 Aufgaben. Im zweiten Teil finden sich vier bis sechs Typ-2-Aufgaben (BIFIE 2013a: 23, Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des Bildungswesens)².

Typ-1-Aufgaben sind Aufgaben, die auf die im Katalog angeführten Grundkompetenzen fokussieren. Bei diesen Aufgaben sind kompetenzorientiert (Grund-)Wissen und (Grund-)Fertigkeiten ohne darüber hinausgehende Eigenständigkeit nachzuweisen.

Typ-2-Aufgaben sind Aufgaben zur Anwendung und Vernetzung der Grundkompetenzen in definierten Kontexten und Anwendungsbereichen. Dabei handelt es sich um umfangreichere kontextbezogene oder auch innermathematische Aufgabenstellungen, im Rahmen derer unterschiedliche Fragestellungen bearbeitet werden müssen und bei deren Lösung operativen Fertigkeiten gegebenenfalls größere Bedeutung zukommt. Eine selbstständige Anwendung von Wissen und Fertigkeiten ist erforderlich. (BIFIE 2013a: 23)

Eine nähere Erläuterung zu den sogenannten Grundkompetenzen folgt im Exkurs zu den Typ-1-Aufgaben auf Seite 7.

Im Falle des Haupttermins 2016 beinhaltete der erste Teil 24 Aufgaben, der zweite Teil umfasste vier Aufgaben. Laut den Grundlagen zur Reifeprüfung des BIFIE (2013a: 33) sind

² Die Handreichung *Die standardisierte schriftliche Reifeprüfung in Mathematik* mit dem Stand von März 2013 ist die ausschlaggebende für den Haupttermin 2015/16 in Mathematik. Auch alle weiteren Publikationen des BIFIE sind in Bezug auf den Haupttermin 2015/16 gewählt. Die neueren Veröffentlichungen betreffen erst die Reifeprüfungen ab dem Haupttermin 2017/18.

„Typ-1-Aufgaben grundsätzlich einer ,0‘- und ,1‘-Beurteilung (richtig/falsch) unterworfen [...], für die ,offeneren‘ Typ-2-Aufgaben [werden] jeweils 0 bis 2 Punkte pro Teilaufgabe vergeben [...].“

Im ersten Teil korrespondiert folglich die maximal mögliche Punktezahl mit der Anzahl der Aufgaben. Im zweiten Teil, der 2016 aus vier Aufgaben bestand, konnten ebenfalls 24 Punkte erreicht werden – die vier Aufgaben wurden in 24 Teilaufgaben unterteilt. Insgesamt besteht dieser Haupttermin also aus 48 Items.

Nachdem bei 48 Items potenzielle Fehler zu analysieren wären und dies 878 mal, müssen die zur Verfügung stehenden Daten eingegrenzt werden. 42.144 einzeln zu untersuchende Aufgaben sprengen den Rahmen der vorliegenden Arbeit. In Absprache mit dem Bundesministerium für Bildung und dem Betreuer wurde beschlossen, die Daten zunächst nach Schultyp einzugrenzen. Dabei fiel die Wahl auf die Langform des Gymnasiums, also die achtjährigen Allgemeinbildenden Höheren Schulen. Dies reduziert die Anzahl der Arbeiten auf 299 vollständige Performanzen.

Wie bereits erwähnt liegen die Performanzen in anonymisierter, digitalisierter Form vor. Dabei sind diese nach Teil 1 und Teil 2 unterteilt. Insgesamt sind 301 Teil 1 Bögen und 299 Teil 2 Bögen eingelangt. Zu jedem Teil-1-Heft existiert ein entsprechendes Teil-2-Heft (bis auf die beiden Teil-1-Hefte, die mehr eingelangt sind). Dies würde nun 14.352 ($299 \cdot 48$) Untersuchungsobjekten entsprechen.

Nach weiterer Beratung mit dem BMB und dem Betreuer wurde der Beschluss gefasst, das Augenmerk auf die Typ-1-Aufgaben, also den ersten Teil, zu legen. Dies hat vor allem den Vorteil, dass sie leichter handzuhaben sind als die „offeneren“ (BIFIE 2013a: 33) Typ-2-Aufgaben. Bei den Typ-1-Aufgaben gibt es, dank der *geschlosseneren* Formate³, (fast) keine Diskussionen, ob ein Punkt zu geben ist oder nicht. Und wenn er nicht gegeben wurde, fängt die Analysearbeit erst an. Die Beschränkung auf die Typ-1-Aufgaben kann als Vereinfachung des Triggers gesehen werden, ob eine Antwort zu einer Aufgabe untersucht werden muss oder nicht.

³ Bei der ssRP Mathematik existieren sechs unterschiedliche Antwortformate: Das offene Antwortformat, das halboffene Antwortformat, Lückentexte, Multiple-Choice Aufgaben, das Zuordnungsformat und das Konstruktionsformat ((BIFIE 2013a: 27-31)). Die geschlossenen Antwortformate *Lückentext*, *Multiple-Choice* und *Zuordnungsformat* kommen vor allem im Teil 1 vor.

Exkurs: Typ-1-Aufgaben

Wie in *Abschnitt 2.1.2* bereits erwähnt handelt es sich bei Typ-1-Aufgaben um Aufgaben, welche die Grundkompetenzen umfassen. Diese werden im ersten Teil der schriftlichen Standardisierten Prüfung Mathematik behandelt (BIFIE 2013a: 23).

Typ-1-Aufgaben fokussieren auf eine Grundkompetenz. Sie werden mit *gelöst* versus *nicht gelöst* verrechnet. Die verschiedenen gebundenen Antwortformate wie Multiple-Choice-Format und Lückentextformat ermöglichen eine exakte Punktevergabe. Zur Vergabe der Punkte bei Aufgaben mit offenem und halboffenem Antwortformat werden Lösungserwartungen bzw. klar formulierte Lösungsschlüssel angegeben. (BIFIE 2013a: 24)

Im Gegensatz dazu stehen Typ-2-Aufgaben, bei denen die Verknüpfung von mehreren Grundkompetenzen im Vordergrund steht. Was soll nun unter dem Begriff der Grundkompetenzen verstanden werden?

In einem mathematischen Kontext handelt es sich bei Grundkompetenzen um „grundlegende mathematische Fähigkeiten bzw. Fertigkeiten [...], die allen Schülerinnen und Schülern längerfristig verfügbar sein sollen“ (BIFIE 2013a: 1). Allgemein gesprochen sind Grundkompetenzen „das Ergebnis von Aushandlungsprozessen“ (BIFIE 2013b: 6). Werner PESCHEK (2011: 213-214) führt dabei das 2009 veröffentlichte Klagenfurter Projekt zur Zentralmatura aus Mathematik (DANGL et al. 2009) an, das Grundkompetenzen als Fähigkeiten beschreibt, „die für das Fach grundlegend sowie gesellschaftlich relevant sind und darüber hinaus längerfristig verfügbar sein sollten sowie leicht („massig“) überprüfbar sein müssen“.

Das Projekt „Standardisierte schriftliche Reifeprüfung aus Mathematik“ (DANGL et al. 2009) bildet auch die Grundlage für das 2013 vom BIFIE veröffentlichte Dokument *Die standardisierte schriftliche Reifeprüfung in Mathematik (BIFIE 2013a)*. DANGL et al. (2009: 13, Hervorhebung im Original) beschreiben die Herkunft der Grundkompetenzen für das Unterrichtsfach Mathematik wie folgt:

Ausgangspunkt und Basis für die Identifizierung von derartigen Grundkompetenzen ist der Lehrplan für den Mathematikunterricht an AHS. Ausgehend von den dort festgelegten Inhaltsbereichen und Inhalten sowie unter Berücksichtigung der angeführten Bildungsziele sind die für die zentrale sRP relevanten mathematischen Grundkompetenzen auszuhandeln und festzulegen. Dabei sind neben fachlichen Aspekten und Zusammenhängen insbesondere auch bildungstheoretische Überlegungen [...] leitend. Das Kernstück der bildungstheoretischen

Orientierung der zentralen sRP bildet eine in einem weiteren Sinne verstandene Lebensvorbereitung, die insbesondere als Befähigung zur *Kommunikation mit Expert(inn)en und mit der Allgemeinheit* gedeutet wird. Eine solche Kommunikationsfähigkeit wird als entscheidendes Orientierungsprinzip für die Auswahl von Inhalten und daran gebundener Kompetenzen gesehen.

So bilden sich vier Inhaltsbereiche in Bezug auf die Grundkompetenzen heraus:

- Algebra und Geometrie (AG)
- Funktionale Abhängigkeiten (FA)
- Analysis (AN)
- Wahrscheinlichkeit und Statistik (WS)

Der Begriff mathematischer Kompetenzen ist allerdings nicht eindimensional. Im *Praxishandbuch Mathematik AHS Oberstufe (Teil 1)* des BIFIE (2013b: 18) wird auf zwei weitere Dimensionen verwiesen: Dies sind zum einen die Handlungsbereiche und zum anderen die Komplexitätsbereiche. Auch diese besitzen wieder spezifische Ausprägungen, welche in Tabelle 1 genannt werden. So ergibt sich im Gesamten folgendes Bild:

<i>Handlungsbereiche</i>	<i>Inhaltsbereiche</i>
H1: Darstellen, Modellbilden	I1: Algebra und Geometrie
H2: Rechnen, Operieren	I2: Funktionale Abhängigkeiten
H3: Interpretieren	I3: Differential- und Integralrechnung ⁴
H4: Argumentieren, Begründen	I4: Wahrscheinlichkeit und Statistik

Komplexitätsbereiche

K1: Einsetzen von Grundkenntnissen und Grundfertigkeiten

K2: Herstellen von Verbindungen

K3: Einsetzen von Reflexionswissen, Reflektieren

Tabelle 1: Modell mathematischer Kompetenzen in der Sekundarstufe 2 (AHS)

Eine konkrete Aufgabe besitzt also eine Handlungs-, eine Inhalts- und eine Komplexitätsdimension, welche als Tripel (H, I, K) dargestellt werden können. Beispielsweise wird die erste Aufgabe des Haupttermins 2016 durch das Tripel (H4, I1, K1) charakterisiert (KRAMER & SATTLBERGER 2016: 69). Bei der ssRP hat sich die bloße Einteilung in

⁴ An Stelle von *Differential- und Integralrechnung* wird der Inhaltsbereich I3 auch als Inhaltsbereich *Analysis* bezeichnet.

Inhaltsbereiche durchgesetzt. So sind auch die Grundkompetenzen, auf denen die Typ-1-Aufgaben basieren, in die vier oben genannten Inhaltsbereiche unterteilt. Typ-1-Aufgaben entsprechen nur den ersten beiden Komplexitätsbereichen.

Bei der ssRP Mathematik wird der erste Teil (also die Typ-1-Aufgaben) in die vier Inhaltsbereiche unterteilt. Diese werden wiederum in Unterbereiche unterteilt. Im Falle des Inhaltsbereichs *Algebra und Geometrie* existieren unter anderem die Unterbereiche *Grundbegriffe der Algebra* und *(Un-)Gleichungen und Gleichungssysteme*, um nur zwei Beispiele zu nennen. Innerhalb dieser werden die einzelnen Grundkompetenzen beschrieben und abgekürzt wie zum Beispiel:

AG 1.1 Wissen über die Zahlenmengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ verständig einsetzen können (BIFIE 2013a: 7).

Eine vollständige Liste der Grundkompetenzen soll hier nicht abgedruckt werden – bei der Bearbeitung der Aufgaben werden die jeweils zugehörigen Definitionen genannt (siehe *Kapitel 4*).

Zusammenfassend kann festgehalten werden, dass Typ-1-Aufgaben genau eine Grundkompetenz abfragen. Diese können mittels dreier Dimensionen charakterisiert werden. Im Falle der schriftlichen standardisierten Reifeprüfung Mathematik werden die Grundkompetenzen in die vier Inhaltsbereiche (*Algebra und Geometrie, Funktionale Abhängigkeiten, Analysis, Wahrscheinlichkeit und Statistik*) aufgeteilt. Die Grundkompetenzen selbst haben ihren Ursprung im Lehrplan und werden unter der Berücksichtigung der Bildungsziele des Curriculums und der bildungstheoretischen Orientierung der ssRP formuliert. Mit dem Abschluss der Ausführungen zu den Typ-1-Aufgaben kann nun der weitere Auswahlprozess beschrieben werden.

Exkurs Ende

Mit der Eingrenzung auf die Typ-1-Aufgaben wurden die Daten schon von 42.144 auf 7.224 ($301 \cdot 24$) zu bearbeitende Items reduziert. Ein letzter Auswahlsschritt folgt noch: Von besonderem Interesse für das Bundesministerium für Bildung sind jene Aufgaben mit geringen Lösungsquoten. Die durchschnittlichen Lösungsquoten der vier Bereiche (Tabelle 2) verhalten sich wie folgt:

Inhaltsbereich	Durchschnittliche Lösungsquote
Algebra und Geometrie	58,6 %
Funktionale Abhängigkeiten	81,6 %
Analysis	65,6 %
Wahrscheinlichkeit und Statistik	62,2 %

Tabelle 2: Durchschnittliche Lösungsquoten Teil 1 (n=301)

In Abbildung 2 werden die einzelnen Aufgaben mit ihren durchschnittlichen Lösungsquoten dargestellt. Abgekürzt werden die Aufgaben jeweils mit X und der entsprechenden Aufgabennummer: So steht beispielsweise X3 für die dritte Aufgabe. Im Detail ergibt sich folgendes Bild:

Lösungsquote Teil 1

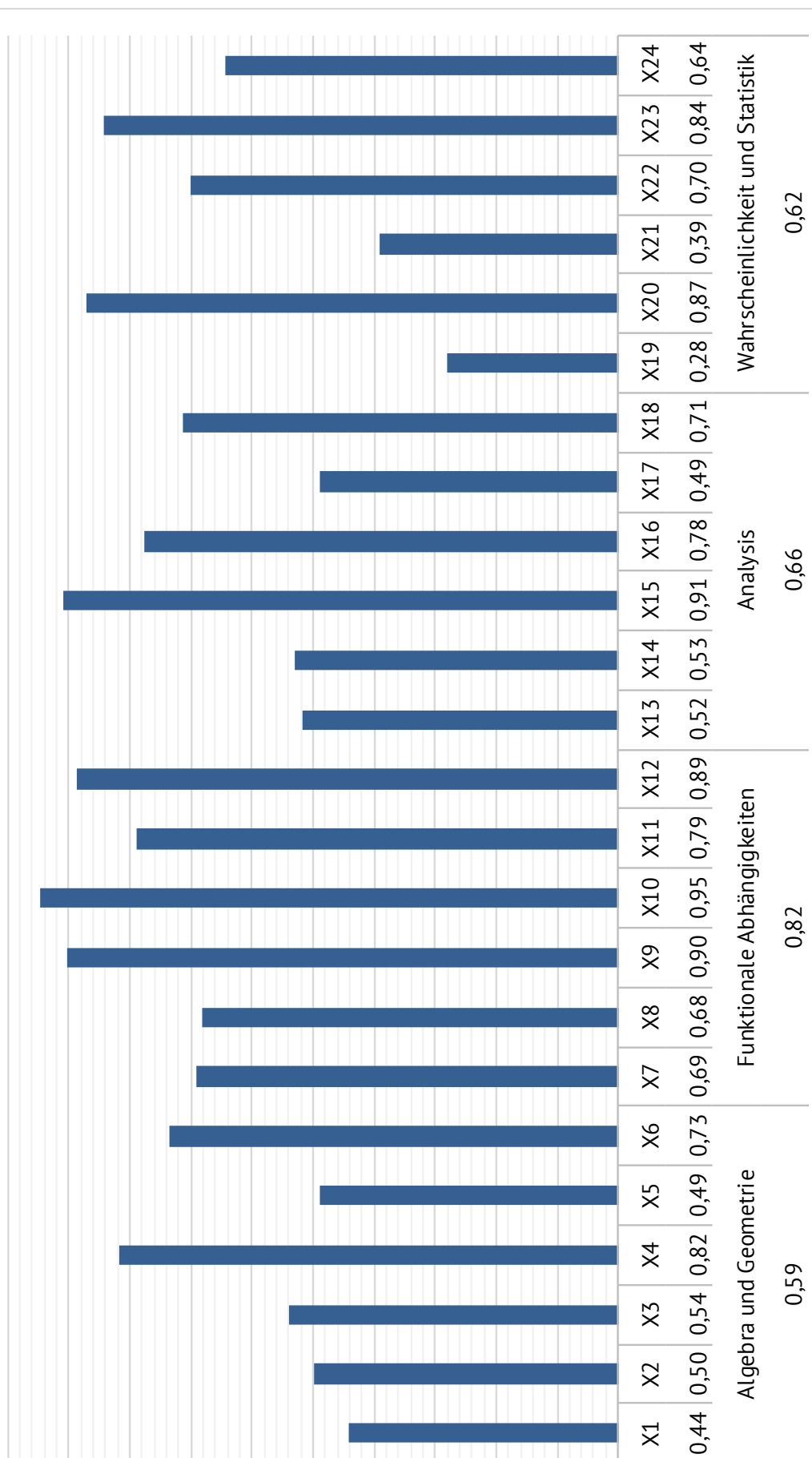


Abbildung 2: Lösungsquote Teil 1 (n=301)

Die beiden Bereiche mit den niedrigsten durchschnittlichen Lösungsquoten sind also *Algebra und Geometrie* sowie *Wahrscheinlichkeit und Statistik*. Auch die drei Items mit den niedrigsten Lösungsquoten (X1, X19 und X21) befinden sich in diesen Bereichen. Das Item X19 konnte beispielsweise nur von 28 Prozent der Kandidatinnen und Kandidaten richtig beantwortet werden. Auf die Bereiche *Algebra und Geometrie* sowie *Wahrscheinlichkeit und Statistik* fällt somit die endgültige Auswahl als die zugrundeliegende Datenmenge für diese Arbeit. Somit werden schlussendlich 3.612 ($301 \cdot 12$) Aufgaben untersucht.

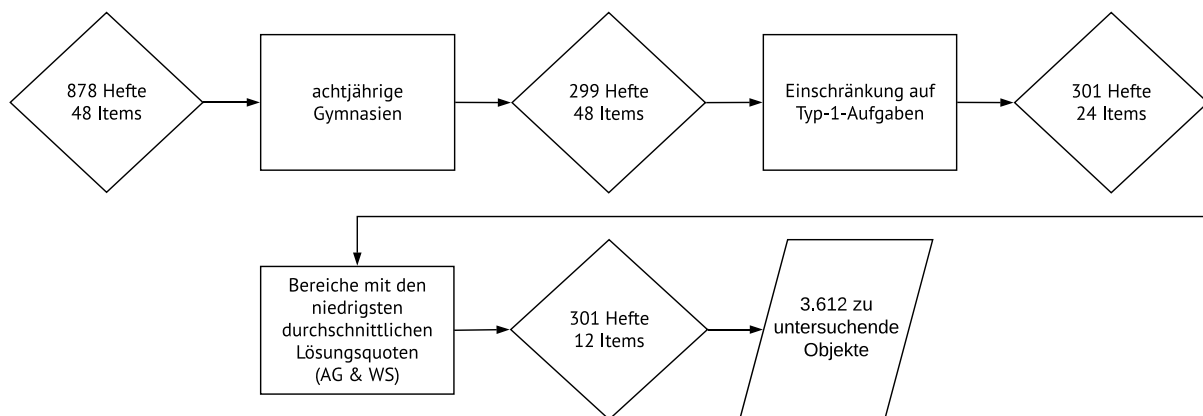


Abbildung 3: Auswahlverfahren

Mit den beschriebenen Einschränkungen (in Abbildung 3 nochmals zusammengefasst) wurde der Umfang von anfangs 42.144 zu untersuchenden Objekten auf 3.612 Aufgaben reduziert. Entsprechend der Anzahl an Heften beziehungsweise Kandidatinnen und Kandidaten liegen die Aufgaben in Form von 301 .pdf-Dateien vor. Details zum Medium der Daten folgen unter *Abschnitt 2.2.1*.

2.2 Prozesse empirischer Forschung

Wie bereits zu Beginn von *Kapitel 2* erwähnt, steht am Anfang jeder Arbeit zunächst die Definition des Problemfeldes (RIESENHUBER 2009: 4). Das Arbeitsfeld hier ist die Zentralmatura (ssRP) Mathematik an AHS 2016 – es stellt sich die Frage mit welchem Ziel diese bearbeitet wird. Am Ende einer wissenschaftlichen Untersuchung muss immer ein Mehrwert entstehen, so auch hier. Schlussendlich sollen Erkenntnisse über gemachte Fehler bei der standardisierten schriftlichen Reifeprüfung Mathematik erlangt werden. Auf dem Weg dahin müssen aber vor allem zuerst die Daten nutzbar gemacht werden. Unter diesem Stern steht die Arbeit. Der Umfang wurde durch eine Auswahl der Daten schon eingeschränkt, trotzdem bleibt die Möglichkeit bestehen, dass nur die Bearbeitung der beiden ausgewählten Teile dieser Arbeit genügt. Auch dann ist ein Ziel erreicht, nämlich die Daten für weitere Projekte aufzubereiten. Bevor explorativ geforscht wird oder Tests durchgeführt werden, steht in der Theoriebildung immer eine deskriptive Forschung (RIESENHUBER 2009: 6). Dies sowie die weiteren Stufen der Theoriebildung sind in Abbildung 4 (RIESENHUBER 2009: 6) zu sehen.

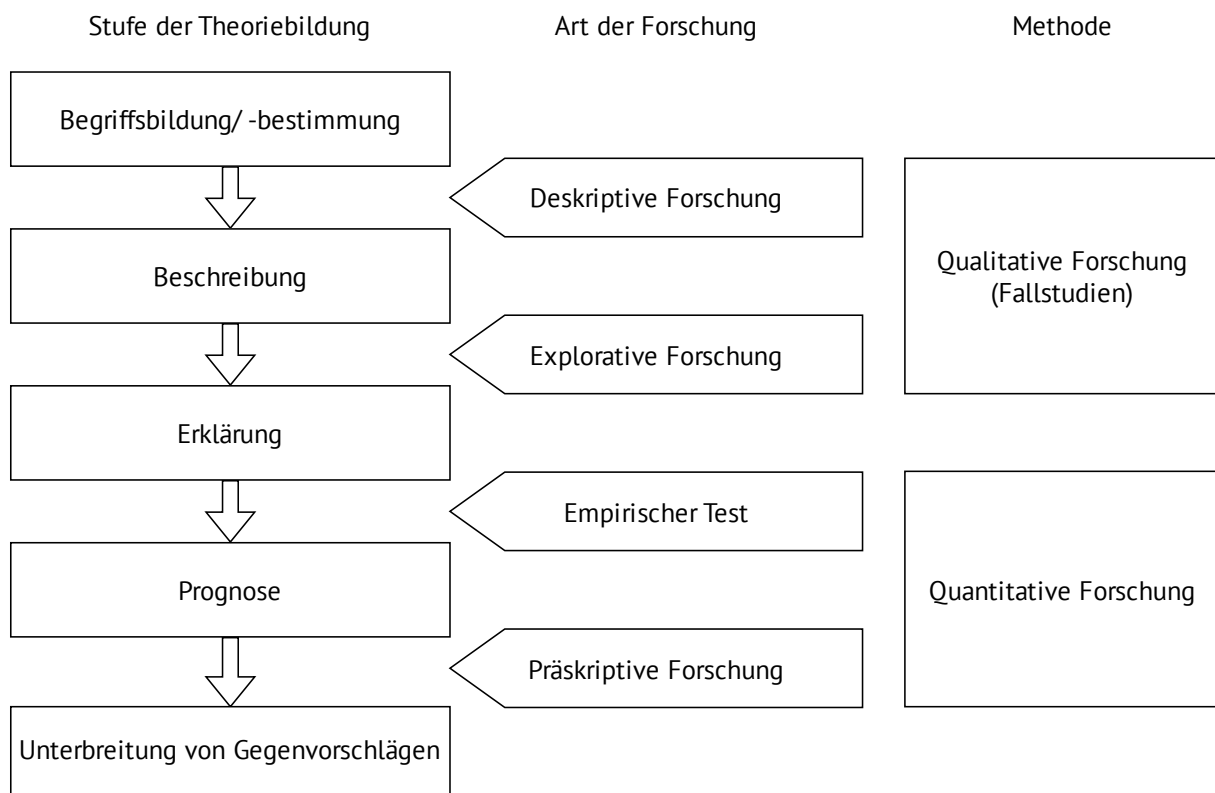


Abbildung 4: Stufen der Theoriebildung, Art der Forschung und Forschungsmethode

Sind sowohl Problemfeld als auch Ziel definiert, folgt als nächstes die Entscheidung zur Forschungsmethode, wobei RIESENHUBER (2009: 4) grundlegend zwischen quantitativer und qualitativer Forschung unterscheidet. Die Forschungsmethode kann dabei etappenweise ermittelt werden: „Eine Forschungsmethode ist das planmäßige und systematische Vorgehen, das zur Gewinnung wissenschaftlicher Erkenntnisse erforderlich ist.“, schreiben HUG & POSCHESCHNIK (2015: 110-111). Um zu diesen Erkenntnissen zu gelangen, werden verschiedene Schritte benötigt, die die Autoren als den „methodischen Dreischritt der empirischen Forschung“ (HUG & POSCHESCHNIK 2015: 81) bezeichnen. Der methodische Dreischritt nennt drei Aspekte methodischen Vorgehens, nämlich Erhebung, Aufbereitung und Auswertung, die helfen können, das Projekt als solches zu strukturieren und die auch ein schrittweises Arbeiten ermöglichen. Für jeden der drei Schritte gibt es, wie in Tabelle 3 (HUG & POSCHESCHNIK 2015: 81) zu sehen, verschiedene Methoden.

Erhebungsmethoden	Aufbereitungsmethoden	Auswertungsmethoden		
		Inhaltsanalyse	Häufigkeitsanalysen	Grounded theory
Recherche	Transkription/ Fixierung			
Dokumentenanalyse	Selegierung	Korrelationsanalyse	Psychoanalyse	Signifikanztest
Beobachtung	Strukturierung	Typenbildung	Clusteranalyse	Diskursanalyse
Befragung	etc.	Konversationsanalyse	Faktorenanalyse	Pfadanalyse
etc.		Metaphernanalyse	Varianzanalyse	etc.

Tabelle 3: Der methodische Dreischritt

Diese drei Schritte werden im Folgenden exemplarisch durchlaufen, um die Daten besser einordnen sowie die in weiterer Folge vorkommenden Methoden besser verorten zu können. Zunächst werden die Erhebungs- und Aufbereitungsmethoden besprochen und die vorliegenden Daten gleich innerhalb dieser Systeme lokalisiert. Bei den Auswertungsmethoden wird sowohl auf qualitative als auch quantitative Forschung eingegangen, wobei auch einige ausgewählte Methoden kurz beschrieben werden. Die Kombination qualitativer und quantitativer Ansätze wird noch angesprochen, bevor in *Kapitel 3* die tatsächlich verwendete(n) Methode(n) vorgestellt wird.

2.2.1 Erhebung

Am Anfang steht immer die Erhebung der Daten – doch welche Möglichkeiten gibt es diese zu erheben? HUG & POSCHESCHNIK (2015: 81) nennen exemplarisch Recherche, Dokumentenanalyse, Beobachtung und Befragung als mögliche Erhebungsmethoden. Die Dokumentenanalyse zeichnet vor allem aus, dass bereits bestehende Daten verwendet werden: „Bei einer Dokumentenanalyse entfällt der Schritt der Datenerhebung bis zu einem gewissen Grad beziehungsweise erschöpft er sich in einem Zusammentragen von existierenden Daten.“ (HUG & POSCHESCHNIK 2015: 82). Auch Thomas BRÜSEMEISTER (2008: 16) sieht in Dokumenten eine Grundlage zur empirischen Forschung. Nicht nur Dokumente in Textform (auch wenn diese den größeren Teil ausmachen) sind zu beachten, auch Bild- und Tonträger fallen in den Bereich der Dokumentenanalyse. Er bemerkt allerdings auch, dass Dokumente nicht die alleinige Quelle einer Untersuchung darstellen müssen, eine Kombination verschiedener Erhebungsmethoden ist ebenso zulässig. Außerdem unterscheidet er zwischen Primär- und Sekundäranalyse. Bei der Primäranalyse wird eine grundlegende Dokumentenanalyse der Originalmaterialien durchgeführt, bei der Sekundäranalyse werden die aufbereiteten Daten beziehungsweise Dokumente genutzt, um weitere Fragestellungen zu beantworten.

Da in dieser Arbeit die unter *Abschnitt 2.1* – bereits vom Bundesministerium für Bildung erhobenen – vorgestellten Daten verwendet werden, entfällt auch hier die Datenerhebung zu einem gewissen Grade. Die Dokumente liegen in schriftlicher Form vor. Bei der Erhebungsmethode handelt es sich folglich um eine Dokumentenanalyse. Da es sich um eingescannte Originale (die Arbeiten der Schülerinnen und Schüler) handelt, muss genauer gesagt jeweils von einem Abbild eines Dokuments in schriftlicher Form gesprochen werden. Diese Feinheit in der Terminologie macht für die weitere Praxis allerdings keinen Unterschied. Es wird eine Primäranalyse durchgeführt, die verfügbaren Rohdaten werden aufbereitet, kategorisiert und so nutzbar gemacht.

Ein Dokument beziehungsweise eine .pdf-Datei umfasst ein komplettes Teil-1-Heft einer Kandidatin oder eines Kandidaten. Dies umfasst das Titelblatt, die Hinweisblätter sowie alle 24 (bearbeiteten) Aufgaben des ersten Teils. Die Hefte sind anonymisiert, weder Name noch Geschlecht werden genannt. Jede Antwort, die auf die Aufgabenbögen geschrieben wurde (und nur diese sind schlussendlich relevant), wurde mit eingescannt. Bei einigen Kandidatinnen und Kandidaten wurden auch noch zusätzliche Blätter mit Nebenrechnungen

(Schmierzettel) digitalisiert, jedoch nicht bei allen. Bei den Arbeiten ohne zusätzliche Blätter wurde allerdings nicht vermerkt, ob keine Schmierzettel existierten oder diese nur nicht eingescannt wurden.

An dieser Stelle wird auch klar, dass das Medium der Daten eine nicht unwesentliche Rolle spielt. Es macht einen großen Unterschied, ob die Hefte im Original vorliegen oder nicht: Liegen sie im Original vor, kann leicht bestätigt oder ausgeschlossen werden, ob Nebenrechnungen gemacht wurden. Liegen sie, wie hier, in digitalisierter Form vor, lässt sich keine Aussage darüber treffen, ob Notizen fehlen oder nicht. Dies beschreibt den ersten von zwei Aspekten von Medien, die HUG & POSCHESCHNIK (2015: 144-145, Hervorhebungen im Original) nennen: „Einmal geht es um die Bedeutung *technischer Hilfsmittel* bei der Gewinnung, Verarbeitung, Speicherung und Übermittlung von Daten; und zum Zweiten geht es um *mediale* Aspekte der Interaktions- und Kommunikationsprozesse im Sinne der Verwendung von Zeichen und ihrer Bedeutung.“

Der Kommunikationsprozess kann bei dieser Art der Daten, genauer gesagt verarbeiteter Daten, als erfolgreich bezeichnet werden, wobei im vorliegenden Fall das beschriebene Papier beziehungsweise das .pdf-File als Medium angesehen wird. Peter ERNST (2011: 32) beschreibt in seiner Einführung zur germanistischen Sprachwissenschaft erfolgreiche Kommunikation:

Damit eine erfolgreiche Kommunikation stattfinden kann, muss es einen Sender geben, der eine Nachricht in einem Medium an einen Empfänger schickt. [...] Dieses einfache Kommunikationsmodell kann nun erweitert werden. Der Sprecher muss eine gedankliche Vorstellung in einen dem Hörer verständlichen Code umsetzen, er muss sie kodieren, und der Empfänger muss sie dekodieren.

Dieses Kommunikationsmodell trifft auch auf die vorliegenden Dokumente zu. Die Kandidatinnen und Kandidaten kodieren ihre Nachricht (die gelösten Aufgaben) mit schriftlichen Zeichen. Durch das Scannen wird zwar ein zusätzlicher Schritt, eine zusätzliche Umwandlung durchgeführt, die Kodierung ändert sich während dieses Prozesses aber nicht. Die schriftlichen Zeichen, die zu Beginn standen, sind schriftliche Zeichen geblieben (anders wäre das zum Beispiel bei einer Verschriftlichung einer mündlichen Prüfung). Geändert hat

sich lediglich das Medium: Die Daten, die vorher noch als sekundäres Medium dargestellt waren, sind jetzt den tertiären Medien zuzurechnen.⁵

HUG & POSCHESCHNIK (2015: 144-145) halten fest, dass die Rolle des Mediums nicht selbstverständlich hingenommen werden darf:

Der zu einer bestimmten historischen Zeit in einer spezifischen Forschungskultur selbstverständliche Gebrauch von Hilfsmitteln täuscht allzu leicht darüber hinweg, dass Forschungsergebnisse in aller Regel in Abhängigkeit von Technologien, Apparaten und Instrumenten zustande gebracht und kritisch überprüft werden können.

Auch hier hat diese Aussage Gültigkeit, ohne die (selbstverständliche) Verwendung von technischen Hilfsmitteln und der Umwandlung der Daten von sekundären zu tertiären Medien, wäre eine Bearbeitung dieser Daten nicht möglich.

2.2.2 Aufbereitung

Für die Aufbereitung von Daten kann zwischen drei verschiedenen Methoden unterschieden werden: Fixierung, Selegierung und Strukturierung (HUG & POSCHESCHNIK 2015: 84). Das Ziel des Aufbereitens ist es, eine erste Struktur in die erhobenen Daten zu bekommen. HUG & POSCHESCHNIK (2015: 134) schreiben:

Aufbereiten meint dabei das Behandeln von erhobenen Daten mit dem Ziel, diese so auszuwählen und zu arrangieren, dass sie den Untersuchungszielen entsprechend, mit dem geringsten Aufwand und dem größtmöglichen Ertrag an Erkenntnissen ausgewertet und dargestellt werden können.

Die Fixierung ist vor allem bei Beobachtungen elementar. Sei es ein mündlich geführtes Interview oder eine visuelle Beobachtung – bevor diese Eindrücke nicht auf irgendeine Art und Weise festgehalten werden, handelt es sich um bloße (subjektive) Erinnerungen. Schon die Nachvollziehbarkeit eines Ereignisses sowie die notwendige Objektivität von Schilderungen bedingen eine Fixierung von Beobachtungen auf ein Medium. Bei einem Interview bieten sich beispielsweise Audio- und/oder Videoaufzeichnungen an,

⁵ Als sekundäre Medien werden Botschaften bezeichnet, die in ihrer Herstellung technische Hilfsmittel verwenden, in der Rezeption aber keine solchen benötigen. Darunter fallen zum Beispiel gedruckte Werke oder alles Geschriebene – in der Herstellung wird Technik benötigt (Schrift, Druck), bei der Rezeption (Lesen) hingegen nicht. Bei den tertiären Medien werden sowohl für die Produktion als auch für die Rezeption auf technische Hilfsmittel zurückgegriffen. Dies ist beispielsweise bei moderner Kommunikation oder Speichermedien (Radio, Fernsehen, digitale Daten) der Fall (HARTMANN 2008: 42-43).

diese können in Folge noch transkribiert werden, was auch einer Form von Fixierung entspricht. Genauso stellen Beobachtungsprotokolle eine Möglichkeit dar, Daten zu fixieren.

In den meisten Fällen werden mehr Daten erhoben, als im Endeffekt verarbeitet werden können. Die Einschränkung auf eine Auswahl, die dann auch tatsächlich ausgewertet wird, nennt man Selegierung (HUG & POSCHESCHNIK 2015: 84). Dabei spielen auch potenzielle Messfehler sowie das Medium der Daten eine Rolle. Da die hier verwendeten Daten auf einer Zufallsstichprobe basieren, muss klar sein, dass unterschiedliche Stichproben in ihrer Auswertung immer (mehr oder weniger) leicht von einer hypothetischen Auswertung der Grundgesamtheit abweichen. Für die vorliegende Stichprobe ist vor allem die Frage nach der Vollständigkeit der Daten relevant. Dies inkludiert zum Beispiel die Frage, ob alle Antworthefte komplett sind. Fehlen Teile eines bestimmten Datensatzes, muss dieser von der Auswertung ausgeschlossen werden. Dies ist bei den vorliegenden Daten nicht der Fall. Alle 301 Hefte wurden vollständig erfasst und können somit in die Auswertung miteinfließen. Darüber hinaus nennen HUG & POSCHESCHNIK (2015: 141) noch Aktualität, Zuverlässigkeit der Datenquelle und technische Aufarbeitung der Daten als wichtige Punkte.

Im Falle der vorliegenden Dokumente wurde die Fixierung bereits durch das BMB erledigt. Wie bereits beschrieben, lagen die Daten von Anfang an in einer schriftlichen Form (handschriftlich) vor, die bereits fixierten Daten wurden lediglich digitalisiert. Die Selegierung erfolgte in Zusammenarbeit mit dem Betreuer und dem BMB, detailliert wurde der Auswahlvorgang bereits unter *Punkt 2.1.2* beschrieben. Zum Zeitpunkt der Bearbeitung handelt es sich bei den vorliegenden Daten um den aktuellsten vollständigen Termin der ssRP Mathematik. Durch das Übermitteln der originalen Arbeiten an das BMB und das Digitalisieren vor Ort darf auch die Zuverlässigkeit der Datenquelle als gegeben erachtet werden.

Nach all diesen Punkten fehlt noch die Strukturierung als dritter Teil der Aufbereitung. Unter die Strukturierung fallen beispielsweise das Sortieren loser Aufzeichnungen oder das Nummerieren von Interviewzeilen. Oft geht die Strukturierung nahtlos in die Auswertung über, da diese häufig mit einer Kategorisierung einhergeht (HUG & POSCHESCHNIK 2015: 84-85). Dank des Formats der Daten (.pdf-Dateien) hält sich der Aufwand bei der Strukturierung in Grenzen. Aus Gründen der Übersichtlichkeit müssen lediglich Typ-1- von Typ-2-Aufgaben

getrennt werden. Diese liegen in weiterer Folge in einer aufsteigenden Nummerierung vor und ermöglichen so eine klare Bezugnahme.

2.2.3 Auswertung

Prinzipiell unterscheiden HUG & POSCHESCHNIK (2015: 85) zwischen qualitativen und quantitativen Auswertungsmethoden. Bei der qualitativen Forschung werden unter anderem die Grounded Theory, die Qualitative Inhaltsanalyse oder die Diskursanalyse genannt. Als Beispiele für quantitative Auswertungsmethoden werden beispielsweise Häufigkeitsanalysen, Zusammenhangs- und Korrelationsanalysen oder auch Pfadanalysen angeführt. Die Frage ob eine bestimmte Thematik nun qualitativ oder quantitativ zu bearbeiten ist, hängt sowohl von der Art der Daten als auch vom Ziel der Forschung ab. HUG & POSCHESCHNIK (2015: 86) sprechen auch von der „Gretchenfrage der Wissenschaft“. Laut Thomas BRÜSEMEISTER (2008: 47) gehen quantitative Methoden meist von einer vorliegenden Hypothese aus, während „es in qualitativen Forschungen nach dem Prinzip der Offenheit darauf [ankommt], sich vom Untersuchungsfeld überraschen zu lassen.“ Dies bedeute aber nicht, ohne jegliche Planung an ein Problemfeld heranzutreten, sondern, dass sich bei qualitativen Untersuchungen Hypothesen oftmals erst im Arbeitsprozess herauskristallisieren.

Felix RIESENHUBER (2009: 6-7) beschreibt qualitative Forschung als Erfassung der „Variabilität der Merkmale des Untersuchungsobjekts durch verbale Beschreibung am Beispiel sorgfältig ausgewählter Einzelfälle.“ Durch die Untersuchung „sorgfältig ausgewählter Einzelfälle“ erlangt ein qualitativer Ansatz detaillierte Einsichten in das zu untersuchende Gebiet. Ziel qualitativer Studien ist es, Hypothesen zu formulieren, „die über die untersuchten Fälle hinaus Gültigkeit haben sollen.“ Auch Thomas BRÜSEMEISTER (2008: 19) nennt Ziele qualitativer Ansätze: „Qualitative Ansätze zielen [...] auf die Entdeckung (Generierung) von Theorieaussagen anhand empirischer Daten.“ Philipp MAYRING (2015: 22-25) nennt sieben Aufgaben qualitativer Analysen:

1. Hypothesenfindung und Theoriebildung
2. Pilotstudien
3. Vertiefungen
4. Einzelfallstudien
5. Prozessanalysen
6. Klassifizierungen
7. Theorie- und Hypothesenprüfung

Zum Punkt Theorie- und Hypothesenprüfung führt MAYRING aus, dass dies in der Regel den quantitativen Ansätzen angerechnet wird, aber eben auch mit qualitativen Methoden möglich sei.

Bei qualitativen Studien müssen oft weniger Daten bearbeitet werden als bei quantitativen Ansätzen (BRÜSEMEISTER 2008: 8). Im Vergleich zu quantitativen Methoden bedarf es also deutlich geringerer Fallzahlen, „um Wissenschaften mit qualitativen Methoden auf neue Spuren zu bringen [...]“ (BRÜSEMEISTER 2008: 19). Qualitative Forschung entdeckt: Oft werden Untersuchungsobjekte in Kategorien eingeteilt, die erst durch die konkrete Arbeit am Forschungsfeld entstehen konnten. Oder wie Thomas BRÜSEMEISTER (2008: 26) es ausdrückt: „Daten rufen dazu auf, einer bestimmten begrifflichen Kategorie zugeordnet (in einem bestimmten Topf getan) zu werden, an welche die ForscherInnen erst durch die Daten selbst aufmerksam wurden.“ In diesem Falle werden also induktive Schlüsse gezogen, es werden Aussagen vom Einzelnen aus über das Ganze getätigt. Allerdings kommt es im Zuge des qualitativen Arbeitens ebenso zu Deduktionen: Beispielsweise ist dies der Fall, wenn aus den Daten zwischenzeitliche Schlüsse gezogen werden und weitere Daten mit diesen Schlüssen verglichen werden. Beim deduktiven Vorgehen wird somit aus einem (vorläufig) gesicherten Ergebnis auf das was folgt geschlossen. Vom Ganzen ausgehend, werden Aussagen über einzelne Elemente getroffen. Der Vergleich zwischen induktivem und deduktivem Vorgehen wird in Abbildung 5 dargestellt. Die sogenannten Zwischenhypothesen, die sich bei der qualitativen Auseinandersetzung mit Daten ergeben, entstehen selbst wiederum oft aus der induktiven Arbeit: So wird am Anfang der Untersuchung nicht aktiv nach solchen Zwischenergebnissen gesucht, während der Analyse ergeben sich dann aber, aufgrund von Auffälligkeiten in den Daten, Zwischenhypothesen. Diesen Arbeitsprozess benennt BRÜSEMEISTER (2008: 26) als „qualitativ-induktiven“ Arbeitsprozess. Im Zuge der qualitativen Forschung werden Zwischenhypothesen durch qualitativ-induktives Vorgehen gefunden, das Arbeiten mit ihnen stellt dann aber einen deduktiven Prozess dar.

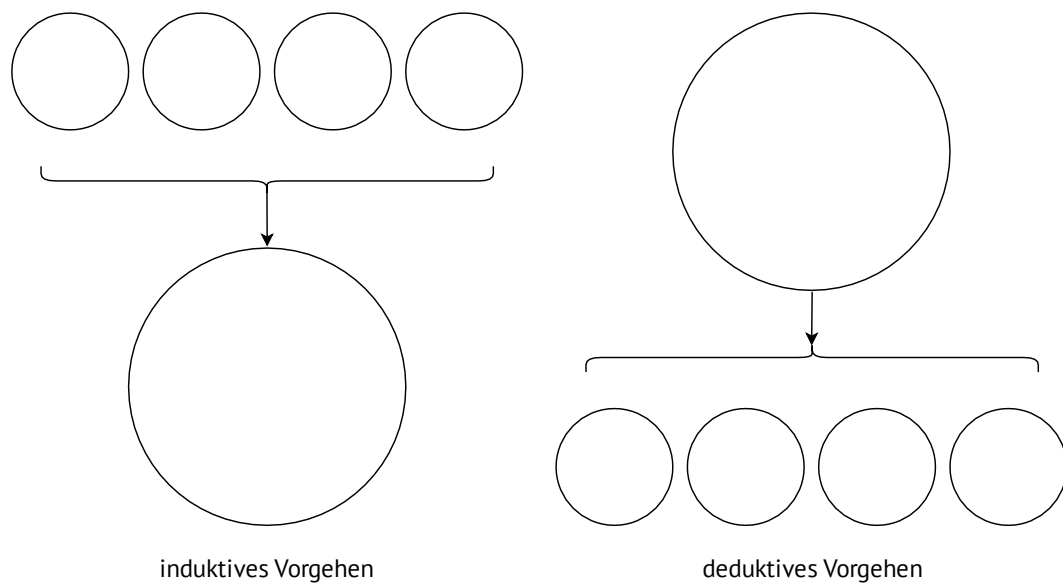


Abbildung 5: Veranschaulichung zu induktivem und deduktivem Vorgehen

Thomas BRÜSEMEISTER (2008: 27-28) führt weiter aus, dass qualitative Forschungsergebnisse nicht nur aus dem Anwenden einer Arbeitsweise (induktiv, deduktiv, ...) entstehen, sondern meist aus einer Kombination dieser:

ForscherInnen [haben] vielleicht eine erste Zwischenhypothese deduktiv an einem zweiten empirischen Datum überprüft; und sie arbeiteten vielleicht qualitativ-induktiv, als sie ihren Untersuchungsgegenstand mit einer Theorie erklärten, die bislang noch nicht auf den Gegenstand angewendet wurde. Schließlich musste die Hauptkategorie, die den Kern des untersuchten Prozesses bezeichnet, vielleicht abduktiv gewonnen werden, weil keine bestehende Theorie die in den Daten benannten Phänomene erklärte. Wesentlich ist, dass qualitative ForscherInnen sich für alle diese Schließungsformen offen halten, um ihre Daten angemessen zu erklären [...].

Allen qualitativen Methoden gemein ist, laut HUG & POSCHESCHNIK (2015: 152), „dass die Auswertung in einem mehrgliedrigen, regelgeleiteten Prozess erfolgt. Man gelangt also Schritt für Schritt von den Daten zur Theorie.“

Es existiert eine unbestimmbare Anzahl von Möglichkeiten Daten qualitativ auszuwerten. HUG & POSCHESCHNIK (2015: 152) nennen beispielsweise die *Grounded Theory*, die *Qualitative Typenbildung*, die *Qualitative Inhaltsanalyse*, *Psychoanalytische Textinterpretation*, die *Diskurs- und Konversationsanalyse* sowie die *Metaphernanalyse* als mögliche Methoden zu einer qualitativen Auswertung. Exemplarisch sollen hier die drei erstgenannten kurz betrachtet werden.

Die Idee hinter der *Grounded Theory* ist die „Entdeckung von Theorie auf der Grundlage von Daten“ (GLASER & STRAUSS 2010: 19). Dabei werden einzelne Elemente des Untersuchungsgegenstands schrittweise zusammengefasst. Zunächst werden die Objekte möglichst konkret und mit Fortdauer der Untersuchung immer abstrakter beschrieben – sie werden kodiert. Dieses Verfahren nennt sich *theoretisches Kodieren* und führt im Endeffekt zu Oberbegriffen, es findet also eine Kategorisierung statt (HUG & POSCHESCHNIK 2015: 155).

Die *Qualitative Typenbildung* stammt aus der empirischen Sozialforschung und zielt wiederum auf eine Kategorisierung ab. Bei diesem Vorgang wird versucht anhand einer Anzahl von Typen Merkmale zu finden, denen sich wiederum einzelne Typen zuordnen lassen (HUG & POSCHESCHNIK 2015: 157). Die Typisierung läuft dabei in drei Schritten ab: Zunächst werden Typisierungsmerkmale herausgefiltert, danach eine Typologie konstruiert. Nachdem eine Typologie entwickelt wurde, werden im dritten Schritt weitere Fälle in das vorhandene System eingeordnet (HUG & POSCHESCHNIK 2015: 157-158). Dabei steht allerdings nicht die Häufigkeit der verschiedenen Typen im Vordergrund, sondern vielmehr die Identifizierung eindeutiger Typen beziehungsweise Kategorien (BRÜSEMEISTER 2008: 30).

Wie bei den beiden vorherigen Vorgehensweisen spielt die Kategorisierung auch bei der *Qualitativen Inhaltsanalyse* eine wichtige Rolle. Das Ziel jeder Inhaltsanalyse stellt laut Philipp MAYRING (2015: 11, Hervorhebung im Original) „die Analyse von Material, das aus irgendeiner Art von *Kommunikation* stammt“ dar. Das Ausgangsmaterial wird dabei in einem ersten Schritt in Sinn-Einheiten eingeteilt. In weiterer Folge werden diese Einheiten immer abstrakteren Kategorien zugeordnet, die die Inhalte beschreiben (HUG & POSCHESCHNIK 2015: 152). Wie auch bei der Qualitativen Typenbildung können ähnliche Einheiten schon beschriebenen Kategorien zugeordnet werden. MAYRING (2015: 51) bezeichnet das Kategoriensystem als den wesentlichen Punkt der Qualitativen Inhaltsanalyse. Kategorien können mittels dreier Theorien entwickelt werden: Der *Definitionstheorie*, dem *Prototypenansatz* und der *Theorie der Entscheidungsgrenzen*. Jede der genannten Theorien besitzt Vor- und Nachteile, und so fordert MAYRING (2015: 48-49) das Anwenden aller drei Theorien, um eindeutig festgelegte Kategorien zu erhalten. Was HUG & POSCHESCHNIK (2015: 152) zur qualitativen Forschung allgemein postulieren, hält auch MAYRING (2015: 50) explizit für die Qualitative Inhaltsanalyse fest: Auch für die Qualitative Inhaltsanalyse stellt das systematische, regelgeleitete Vorgehen einen wesentlichen Punkt dar. Bevor die Qualitative Inhaltsanalyse noch detaillierter beschrieben werden soll, folgen noch einige Punkte zur

quantitativen Forschung, um schlussendlich der Frage nachgehen zu können, ob sich in der, von HUG & POSCHESCHNIK (2015: 85) beschriebenen, Trennung von qualitativer und quantitativer Forschung nicht doch verbindende Elemente finden lassen.

Gleich zu Beginn seiner Ausführungen zur *quantitativen Forschung* schreibt Felix RIESENHUBER (2009: 7): „Im Gegensatz zur verbalen Beschreibung qualitativer Forschung erfasst quantitative Forschung die Variabilität eines Merkmals über die definierte Zuordnung von Zahlenwerten.“ Dabei werden bestimmte Merkmale definiert und erhoben wie oft die dazugehörigen Merkmalsausprägungen vorkommen. Somit lassen sich mithilfe quantitativer Methoden größere Datenmengen signifikant einfacher untersuchen, als dies mit qualitativen Methoden der Fall wäre. RIESENHUBER (2009: 7) führt weiter aus, dass durch die Möglichkeit mehr Daten zu bearbeiten, „die Repräsentativität der Ergebnisse für die Grundgesamtheit“ höher einzustufen ist. Als Ziel quantitativer Forschung gibt er demnach die Überprüfung von Forschungshypothesen an. Diese werden schlussendlich empirisch widerlegt oder bestätigt und somit in statistische Hypothesen umgewandelt. So beschreibt auch Thomas BRÜSEMEISTER (2008: 28) quantitative Forschung: In der Regel sind Forscherinnen und Forscher schon im Besitz einer Theorie, „aus der Variablen erstellt werden, die man für eine empirische Erhebung operationalisiert.“ Es wird demnach deduktiv gearbeitet.

HUG & POSCHESCHNIK (2015: 111) definieren die quantitative empirische Sozialforschung als Vorgang, Daten „auf möglichst objektivierte Weise zu beschreiben, grundlegende Gesetzmäßigkeiten und Zusammenhänge zu entdecken und überprüfbar zu machen“. Dies alles auf der Grundlage von zahlenmäßigen Angaben (Hervorhebungen im Original):

Quantitative Methoden in der empirischen Sozialforschung umfassen alle Praktiken zur *numerischen*, d. h. *zahlenmäßigen*, also bspw. auf Prozentangaben fußenden Darstellung *empirischer* (auf *Beobachtungen der Realität* gründender) Sachverhalte. [...] Der quantitative Ansatz beinhaltet sowohl die Stichprobenauswahl, die Datenerhebung (*»Messung«*) selbst als auch die Auswertung (*»Analyse«*) des Datenmaterials.

Grundlage für quantitative Forschung sind Daten, „die einen *quantitativen* (numerischen) Vergleich mit anderen Daten ermöglichen.“ (HUG & POSCHESCHNIK 2015: 139-140, Hervorhebungen im Original). Dabei können die Daten schon aus früheren Untersuchungen vorliegen oder neu erhoben werden. Wichtig ist, dass über allem ein konkreter Forschungsplan steht. Werden die Daten durch den Einsatz von standardisierten

Messwerkzeugen quantifiziert, wird vom Messen gesprochen (HUG & POSCHESCHNIK 2015: 139-140). BAUR & BLASIUS (2014b: 57) weisen ausdrücklich darauf hin, „dass schon bei der Datenerhebung die Datenauswertung berücksichtigt werden muss.“ So sei bei der Arbeit mit Skalen beispielsweise unabdingbar, mehrere Items zum selben Thema zu erheben. Ist die Erhebung abgeschlossen, folgt die Frage, wie Daten quantitativ ausgewertet werden können.

Dazu nennen HUG & POSCHESCHNIK (2015: 182-183) die *deskriptive Statistik* sowie die *schließende Statistik (Inferenzstatistik)*. Die deskriptive Statistik bedient sich vor allem Kennzahlen und verschiedenen Darstellungsformen, um Aussagen treffen zu können. Zu den wichtigsten Kennzahlen gehören beispielsweise der Modus, der Median, das arithmetische Mittel, die Varianz und Standardabweichung und der Korrelationskoeffizient. Als Beispiele für verschiedene Darstellungsformen nennen HUG & POSCHESCHNIK die tabellarische Darstellung von Daten. Dies kann in Form von Häufigkeitstabellen oder auch mithilfe von Histogrammen geschehen. Die Inferenzstatistik fragt nach dem Zusammenhang zwischen wissenschaftlichen Ergebnissen und dem Zufall, der jeder Stichprobe innewohnt. HUG & POSCHESCHNIK (2015: 183) schreiben: „Entsprechend geht es auch darum, inwieweit eine Stichprobe Rückschlüsse auf die jeweilige Grundgesamtheit, aus der die jeweilige Stichprobe ‚gezogen‘ wurde, zulässt; d.h.: inwieweit sie repräsentativ ist.“ Das wichtigste Werkzeug dabei ist der Hypothesentest, der Aussagen darüber trifft ob eine *alte* Hypothese im Vergleich zu einer *neueren* mit einer bestimmten (Irrtums-)Wahrscheinlichkeit nicht abgelehnt werden kann.

Im *Handbuch Methoden der empirischen Sozialforschung* von BAUR & BLASIUS (2014a) werden verschiedenste quantitative Auswertungsmethoden genannt. Beispielhaft sollen hier nur noch das *Skalierungsverfahren* und die *Längsschnittanalyse* kurz beschrieben werden. „Unter Skalierung versteht man, wie ein oder mehrere Beobachtungsmerkmal(e) einem oder mehreren Skalenwert(en) zugeschrieben wird (werden). Dabei sind die Beobachtungsmerkmale *manifest*, sie werden direkt erhoben [...]“, so Jörg BLASIUS (2014: 1051, Hervorhebung im Original), der gleich ein Beispiel anfügt. Die Antwort auf die Frage: „Glauben Sie an Gott?“, wird mithilfe eines Skalierungsverfahrens individuell zusammengefasst und skaliert. In der Gesamtheit aller Befragten ergibt sich so eine Skala, die eine Aussage über die Religiosität der Befragten zulässt.

Eine Längsschnittanalyse untersucht ein bestimmtes Merkmal über einen längeren Zeitraum. MIKA & STEGMANN (2014: 1077) beschreiben das wie folgt: „Längsschnittdaten liegen vor, wenn zu einer Befragungseinheit von Personen [...] Daten zur Verfügung stehen, die nicht nur die Situation zu einem einzigen Messezeitpunkt, sondern über einen längeren Zeitraum erfassen.“ Die Dauer kann dabei je nach Forschungsfrage variieren. Mithilfe von Längsschnittdaten kann die Veränderung eines Merkmals im Laufe der Zeit beschrieben werden, dadurch kann zum Beispiel das Absinken oder Ansteigen der Arbeitslosigkeit beobachtet werden (MIKA & STEGMANN 2014: 1077).

Fasst man nun die Merkmale qualitativer und quantitativer Forschung zusammen, so ergibt sich ein Überblick wie in Tabelle 4, die bei HUG & POSCHESCHNIK (2015: 112) zu finden ist:

Merkmale <i>quantitativer</i> Forschung	Merkmale <i>qualitativer</i> Forschung
• systematische, standardisierte Messungen von empirischen (auf realen Fakten gründenden) Sachverhalten • Verfahren zum Testen von Hypothesen • meistens Untersuchung großer Fallzahlen / großer Stichproben • Objektive Messung und Quantifizierung von Sachverhalten • Messung zählbarer Eigenschaften • Auswertung durch statistische Instrumente, Analyse statistischer Zusammenhänge	• dienen oft der Entwicklung neuer Hypothesen (oft auf relativ neuen Forschungsgebieten) • Relativ offenes und flexibles Vorgehen (oft ist nur ein grober thematischer Leitfaden gegeben) • kleine Zahl von Untersuchungspersonen, dafür tiefer gehende Betrachtungen und Einzelfallanalysen • meistens kein Anspruch auf Repräsentativität • keine statistische Auswertung • auf das Verstehen von Sinn (etwa von persönlichen Handlungsmotiven, Absichten) bezogen • subjektive Faktoren stehen im Vordergrund, die nicht gemessen, sondern interpretiert werden

Tabelle 4: Merkmale quantitativer und qualitativer Forschung

So gegensätzlich die beiden Zugangsweisen nun erscheinen mögen, so bedingen sie einander doch. Dies beschreibt auch Thomas BRÜSEMEISTER (2008: 35): „Qualitative Forschungsergebnisse sind oft Ausgangspunkt für weiterführende quantitative Untersuchungen. Wurden zuerst repräsentative Merkmale beschrieben, handelt es sich dann um eine quantitative Einordnung der Phänomene.“ Die sture Anwendung von nur einem der beiden Ansätze gehört der Vergangenheit an, vielmehr wird mit Kombinationen von qualitativen und quantitativen Ansätzen gearbeitet. Diese Herangehensweise wird *Mixed Methods* genannt und wird laut Udo KELLE (2008: 15) schon seit den 1980er Jahren immer häufiger verwendet. „Die Methoden werden dabei parallel oder sukzessive eingesetzt und die Resultate aufeinander bezogen. Die Ergebnisse der Methodentypen können sich dann gegenseitig validieren oder einander ergänzen.“ (HUG & POSCHESCHNIK 2015: 90).

Eine eindeutige Zuordnung eines Forschungsprojekts zu einer der beiden Auswertungsmethoden ist oft nicht möglich und auch keineswegs zwingend, wie „Arbeiten von Vertretern der qualitativen Forschungstradition [...] zeigen“ (KELLE 2008: 53). Auch HUG & POSCHESCHNIK (2015: 90) machen darauf aufmerksam, dass in der Praxis der empirischen Forschung der Ansatz der Mixed Methods immer öfter zum Einsatz kommt. Sind Daten einmal erhoben und die Auswertung steht an, so kann dies auf „qualitative [...] oder eben quantitative Art und Weise geschehen, wobei sich diese beiden Ansätze aber keinesfalls gegenseitig ausschließen.“ (HUG & POSCHESCHNIK 2015: 164). So fließen beispielsweise in qualitativen Analysen oft auch quantitative Aspekte mit ein: In vielen Fällen werden im Zuge qualitativer Untersuchungen Kategorien gebildet, woraufhin sich die „Häufigkeiten der Ausprägungen“ untersuchen lassen (MAYRING 2015: 18-19).

Und so wird im folgenden *Kapitel 3* nicht auf die Methode, sondern auf die Methoden eingegangen, die zur Auswertung der Daten herangezogen werden. Denn für die am Beginn des Abschnitts erwähnte „Gretchenfrage der Wissenschaft“ finden HUG & POSCHESCHNIK (2015: 86) eine passende Replik:

Wir möchten die strenge Polarisierung zwischen qualitativ und quantitativ vermeiden und lieber einen pragmatischen Mittelweg einschlagen, denn bei nüchterner Betrachtung haben eigentlich alle Forschungsmethoden gewisse Stärken und gewisse Schwächen. Jede Methode – egal ob qualitativ oder quantitativ – ist für bestimmte Forschungsgegenstände und Forschungsfragen besser und für andere schlechter geeignet.

3 Methodik

Nachdem geklärt wurde woher die Daten stammen, wie diese aussehen beziehungsweise aufbereitet wurden und ein Überblick über die verschiedenen Ansätze zur empirischen Forschung geschaffen wurde, ist es an der Zeit, die Vorgehensweise, die zur Bearbeitung der vorliegenden Daten gewählt wurde, näher zu beschreiben und zu begründen. Wie am Ende des vorherigen *Abschnittes 2.2.3* klar wurde, ist es oft nicht möglich eine Problemstellung einem qualitativen oder einem quantitativen Ansatz zuzuordnen. Nicht nur das, es ist auch gar nicht notwendig.

Um vor Augen zu führen, dass auch diese Arbeit nicht eindeutig einem der beiden Ansätze zugeordnet werden kann, genügt ein Vergleich einiger Merkmale aus Tabelle 4 in *Abschnitt 2.2.3* von HUG & POSCHESCHNIK (2015: 112). Dort wird als eines der grundlegendsten distinktiven Merkmale die Anzahl der untersuchten Objekte genannt. Große Fallzahlen werden eher der quantitativen Forschung zugeschrieben, eine geringere Anzahl an Untersuchungsobjekten deutet eher auf einen qualitativen Ansatz hin. Ein kleiner Rückblick zu *Abschnitt 2.1.2*: Zur Verfügung stehen 878 Hefte, die jeweils 48 Aufgaben beinhalten. So könnten theoretisch 42.144 Objekte untersucht werden. Nach mehreren Schritten wurde die Auswahl auf 301 Hefte und zwölf Aufgaben beschränkt, also 3.612 zu untersuchende Objekte. Insgesamt entspricht das einem Anteil von 8,57 Prozent aller zur Verfügung stehender Items. Ist das nun ausreichend um sich für eine quantitative Untersuchung zu qualifizieren oder muss hier – aufgrund der Anzahl der untersuchten Objekte – von einer qualitativen Studie ausgegangen werden? Diese Frage kann so nicht beantwortet werden. Auch wenn der Untersuchungsgegenstand auf verschiedenen Ebenen betrachtet wird, ergibt sich kein eindeutiges Bild. Auf der einen Ebene werden zwölf Aufgaben behandelt, auf einer Ebene darunter entsprechen diese zwölf Aufgaben aber wiederum 3.612 Objekten – auch so ergeben sich vollkommen unterschiedliche Anzahlen.

Neben der Anzahl soll noch ein anderes Merkmal betrachtet werden: Auf der einen Seite spielen Messungen eine sehr wichtige Rolle in dieser Untersuchung und mit diesen auch die Quantifizierung, was auf ein quantitatives Vorgehen schließen lassen würde. Die Frage die sich nun stellt, ist aber: Was soll quantifiziert werden? Stehen am Ende der Arbeit 3.612 Kategorien mit jeweils einer Entsprechung, ist die Aussagekraft doch in Frage zu ziehen. Zunächst müssen also Kategorien entwickelt werden, die ähnliche Ausprägungen

zusammenfassen. Dazu bedarf es „tiefer gehende[r] Betrachtungen“ (HUG & POSCHESCHNIK 2015: 112) sowohl der Aufgaben, als auch der Ausarbeitungen dieser. Dieser Vorgang ist qualitativ einzustufen. Das eine bedarf also auch hier des anderen, qualitativ und quantitativ gehen Hand in Hand. Zunächst müssen Kategorien gefunden werden, dies erfolgt nach der Vorgehensweise der induktiven Kategorienbildung, wie sie Philipp MAYRING (2015: 85-90) in seinen Ausführungen zur Qualitativen Inhaltsanalyse beschreibt. Erst auf der Grundlage qualitativer Untersuchungen können in weiterer Folge quantitative Aussagen getroffen werden, wie auch MAYRING (2015: 21) schreibt. Diese quantitativen Aussagen werden auf der Basis einer Häufigkeitsanalyse erfolgen.

3.1 Induktive Kategorienbildung

Kategorienbildung oder allgemeiner gesprochen Klassifizierung hat zunächst den Sinn erhobene Daten zu ordnen. Nach welchen Gesichtspunkten die Daten dabei geordnet werden, hängt von den empirischen und theoretischen Umständen der Forschung ab (MAYRING 2015: 24). Als „deskriptive Analyseraster“ bezeichnet Bernd REINHOFER (2008: 125) Kategorien. In ihrer beschreibenden Funktion müssen Kategorien deshalb „abstrakter [...] sein als das Material, das sie durch Überbegriffe ordnen bzw. ordnen sollen.“ Diese deskriptiven Systeme erklärt MAYRING (2016: 100):

Der Grundgedanke bei der Konstruktion deskriptiver Systeme ist dabei, dass sie immer abstrakter als das Material sind, da sie dieses ja ordnen sollen. Sie stellen Verallgemeinerungen dar, hervorgegangen aus der Ebene konkreter empirischer Tatbestände; sie sind somit theoretische Aussagen.

Michael WALDMANN (2016: 378) beschreibt Kategorien als Mittel „um Gruppen von Objekten oder Ereignissen [...] aufgrund ihrer Gemeinsamkeiten zusammenzufassen“. Die Frage nach dem Nutzen von Kategorien beantwortet WALDMANN (2016: 378) gleich selbst: „Eine der wichtigsten Funktionen von Kategorien besteht darin, dass sie uns ermöglichen, neue Erfahrungen mit bereits vorhandenem Wissen in Kontakt zu bringen.“ Darüber hinaus würden Kategorien Lernprozesse ermöglichen, Inferenzen unterstützen und stellen oft auch Grundlagen für Erklärungen dar, um nur einige Beispiele zu nennen. WALDMANN (2016: 379) schreibt weiter:

Zusammenfassend dienen Kategorien einer Vielzahl ganz unterschiedlicher Funktionen, die in mannigfacher Weise interagieren können. Kategorien sind nicht nur die Basis von Klassifikation, sie spielen auch eine zentrale Rolle in einer Reihe wichtiger kognitiver Prozesse. Neben ihren

vielfältigen adaptiven Funktionen können Kategorien allerdings auch zu Verzerrungen der Übereinstimmungen und Unterschiede der von ihnen zusammengefassten Objekte führen.

Damit weist er auch auf mögliche Probleme hin, die die Arbeit mit Kategorien mit sich bringen kann. Doch wie können solche Kategorien gebildet werden? Und wie kann festgelegt werden welche Objekte welchen Kategorien zugehörig sind?

WALDMANN (2016: 380-393) bezieht sich dazu vor allem auf *ähnlichkeitsbasierte Ansätze* und unterscheidet dabei zwischen fünf Ansätzen: der klassischen Sicht, der Prototypensicht, der Exemplarsicht, der Theorie der Entscheidungsgrenzen und der Theorien multipler Systeme. Eleanor ROSCH (1978: 32, in Übersetzung von WALDMANN 2016: 380) schreibt zu ähnlichkeitsbasierten Ansätzen: „Dieses Verfahren erzeugt Klassen, bei denen die Ähnlichkeit innerhalb der Kategoriengrenzen maximiert und die Ähnlichkeit zwischen Kategorien minimiert wird.“ MAYRING (2015: 48) unterscheidet im Wesentlichen zwischen drei Kategorisierungstheorien: der *Definitionstheorie* (entspricht der klassischen Sicht bei WALDMANN), dem *Prototypenansatz* (Prototypensicht bei WALDMANN) und der *Theorie der Entscheidungsgrenzen*. Diese drei von MAYRING genannte Theorien werden etwas näher betrachtet.

„Die Definitionstheorie [...] formuliert explizit notwendige und hinreichende Bedingungen für die Zugehörigkeit eines Elements zu einer Kategorie.“, so MAYRING (2015: 48). Sind diese Regeln einmal gefunden sollte theoretisch jedes Objekt eindeutig einer Kategorie zugeordnet werden können. Darüber hinaus muss gelten, dass alle Entsprechungen einer Kategorie gleich gute Beispiele für diese Kategorie sind (WALDMANN 2016: 380). Genau da liegt aber auch ein Problem der Definitionstheorie: WALDMANN (2016: 381) schreibt weiter, „dass Kategorien nicht diskret sind, sondern dass verschiedene Exemplare der Kategorie als mehr oder weniger stark der Kategorie zugehörig empfunden werden.“ Beispielsweise würde ein Tisch als typischeres Mitglied der Kategorie Möbelstück gelten als ein Teppich. Als grundlegendes Problem der Definitionstheorie sieht MAYRING (2015: 48), „dass diese Definitionen nie ausreichend für alle Fälle sein können und auch im Prozess des Lernens einer Kategorie zu komplex werden.“

Auch die Prototypentheorie kommt zum Schluss, dass nicht alle Entsprechungen für eine Kategorie gleich typisch sind (WALDMANN 2016: 380). „Gemäß dieser Sicht werden Kategorien als Bündel von Merkmalen repräsentiert, die typisch oder charakteristisch für die

Kategorie sind, aber nicht unbedingt notwendig oder hinreichend.“, so WALDMANN (2016: 381) weiter. Damit kann ein Objekt einer Kategorie mehr oder weniger angehören. Ein Prototyp bündelt dabei charakteristische Merkmale einer Kategorie, ohne dass dieser notwendigerweise selbst alle Merkmale aufweisen muss. Neue Objekte werden Kategorien zugeordnet indem sie mit den, für diese Kategorie typischen, Prototypen verglichen werden (WALDMANN 2016: 381-382). Problematisch beim Prototypenansatz der Kategorisierung ist, laut MAYRING (2015: 48), die Abgrenzung zu anderen Kategorien: Beispielsweise sei die Abgrenzung zwischen Baum und Busch, je nach Prototypenwahl, nicht immer eindeutig.

Dieser Uneindeutigkeit Abhilfe schaffen soll der dritte Ansatz der von MAYRING (2015: 48) genannt wird: Die Theorie der Entscheidungsgrenzen soll durch das Festlegen von Abgrenzungsregeln eine klare Trennung zwischen verschiedenen Kategorien schaffen. „Kategorisierung besteht dieser Theorie zufolge also nicht in einem Ähnlichkeitsvergleich [...] mit einem Prototyp [...], sondern darin, dass der Beobachter entscheidet, in welche Region der jeweilige Stimulus fällt und die dementsprechende Kategorienentscheidung trifft.“ (WALDMANN 2016: 388). Dabei sind die Entscheidungsgrenzen klar von den Regeln der Definitionstheorie abzugrenzen: Während bei der Definitionstheorie jedes Objekt einer Kategorie ein gleich gutes Beispiel für diese Kategorie ist, geht die Theorie der Entscheidungsgrenzen davon aus, dass Objekte, die sich näher an den Entscheidungsgrenzen befinden, schwieriger zuzuordnen sind, als Objekte die inmitten zweier Grenzen liegen (WALDMANN 2016: 388). Kritik gibt es auch an diesem Ansatz, wie MAYRING (2015: 48) klar macht: Auch dieser Ansatz sei alleinig nicht dazu geeignet, Kategorisierungsprozesse zu beschreiben.

Für jede der drei genannten Theorien gibt es also gute Gründe sowie auch berechtigte Kritik. Deshalb kommt MAYRING (2015: 48-49) zum Schluss, „dass alle drei Ansätze (Definitionen, Prototypen, Abgrenzungsregeln) Erklärungskraft besitzen und wohl im Alltag ineinandergreifen. [...] Wenn Kategorien eindeutig festgelegt werden sollen, bedarf es also aller drei Definitionsprozesse.“ Die Kombination verschiedener Ansätze nennt WALDMANN (2016: 389-390) die Theorie multipler Systeme. Alle drei vorgestellten Theorien und somit auch die Theorie multipler Systeme basieren auf Ähnlichkeit. „Diejenigen Objekte und Ereignisse werden zusammengefasst, die sich untereinander ähnlich sind, und Kategorisierung beruht darauf, dass man die Ähnlichkeit zwischen neuen Objekten und den bereits vorhanden Kategorien bestimmt.“ (WALDMANN 2016: 393). Allerdings wird dabei

verneint, dass Kontext und Vorwissen einen Einfluss auf die Kategorisierung haben können. Laut WALDMANN (2016: 396) ist Ähnlichkeit dabei „nicht nur eine mögliche Ursache von Kategorien, es gibt auch Belege, die zeigen, dass sie ein Produkt von Kategorisierungsprozessen sein kann.“ Mit dem Hinweis, dass das Prinzip der Ähnlichkeit nichtsdestotrotz einen elementaren Teil von Kategorisierungsprozessen darstellt, schließt WALDMANN (2016: 396) seine Ausführungen zu den Kategorisierungstheorien.

Innerhalb der Qualitativen Inhaltsanalyse unterscheiden HUG & POSCHESCHNIK (2015: 153) zwischen zwei verschiedenen Arten mit Kategorien zu arbeiten: „Erstens eine deduktive Kategorienanwendung und zweitens eine induktive Kategorienbildung.“ Eine deduktive Vorgehensweise bedingt schon vor dem Beginn der Auswertung eine Theorie beziehungsweise ein Kategoriensystem, das auf das vorliegende Material angewendet werden kann. Bei einer induktiven Kategorienbildung werden – genau umgekehrt – aus den Daten Kategorien erarbeitet, die dann mit der weitergehenden Durchsicht des Materials schrittweise angepasst werden. Deduktiver und induktiver Ansatz sind in diesem Zusammenhang bloß zwei Seiten einer Medaille, laut HUG & POSCHESCHNIK (2015: 153) werden meistens beide Strategien kombiniert. Auch Uwe FLICK (1995: 165) weist darauf hin, dass keine Forscherin beziehungsweise kein Forscher weder „völlig vorbehaltlos aus den Daten heraus, noch völlig bruchlos aufgrund seines theoretischen Hintergrundes kategorisieren“ wird.

Wie bereits in *Kapitel 2* erwähnt, gilt es die getroffene Auswahl aus der Stichprobe nutzbar zu machen. Aufgrund der Aktualität des Materials gibt es noch keine vorbereitenden Studien zu diesen Daten. Die deskriptive Aufgabe kommt also dieser Arbeit zu und soll mit Hilfe von induktiver Kategorienbildung erfolgen. MAYRING (2015: 85) schreibt: „Eine induktive Kategoriendefinition [...] leitet die Kategorien direkt aus dem Material in einem Verallgemeinerungsprozess ab, ohne sich auf vorab formuliert Theorienkonzepte zu beziehen.“ Dabei wird ein sogenannter bottom-up-Ansatz verfolgt (REINHOFFER 2008: 125). Dem entgegen steht der top-down-Ansatz. Im Wesentlichen beschreiben die beiden Ansätze die induktive und deduktive Strategie beim Kategorisierungsprozess (HUG & POSCHESCHNIK 2015: 56). Laut MAYRING (2015: 86) strebt die induktive Kategorienbildung beziehungsweise der bottom-up-Ansatz der Kategorisierung „nach einer möglichst naturalistischen, gegenstandsnahen Abbildung des Materials ohne Verzerrungen durch Vorannahmen des Forschers, einer Erfassung des Gegenstandes in der Sprache des Materials.“

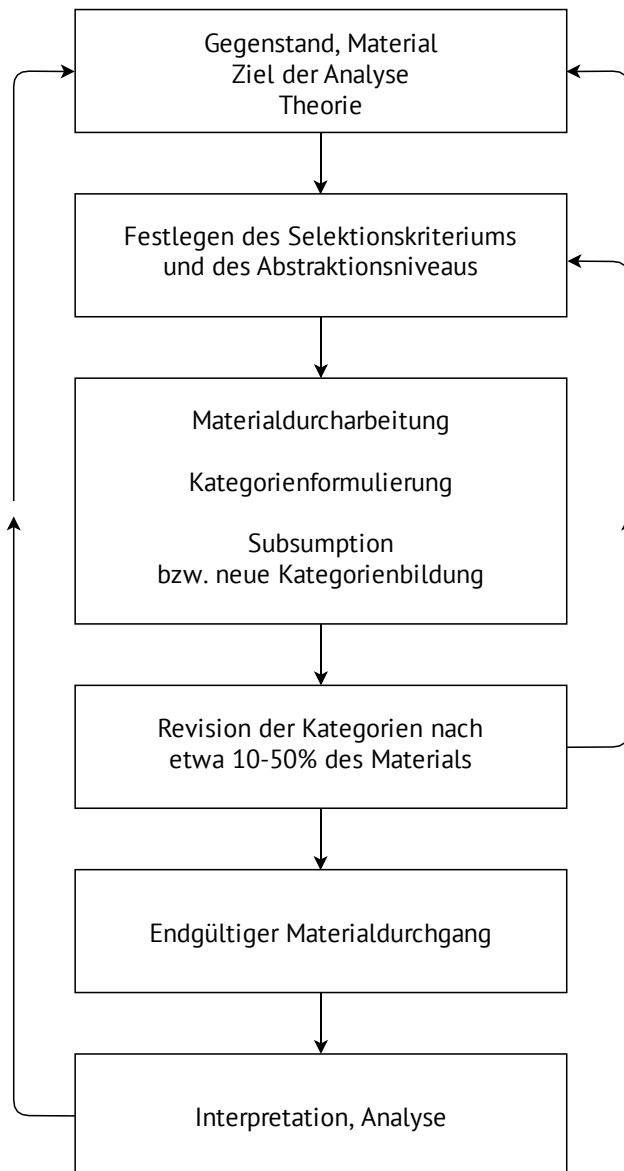


Abbildung 6: Prozessmodell induktiver Kategorienbildung

Bevor die induktive Arbeit beginnen kann, wird zuerst das Thema der Kategorienbildung fixiert. Dafür muss ein Selektionskriterium bestimmt werden, dieses legt fest, „welches Material Ausgangspunkt der Kategoriendefinition sein soll. Dadurch wird Unwesentliches, Ausschmückendes, vom Thema Abweichendes ausgeschlossen.“ (MAYRING 2015: 86-87). Im vorliegenden Fall werden diejenigen Aufgaben untersucht, die fehlerhaft bearbeitet wurden. Die Entscheidung, ob eine Aufgabe richtig gelöst wurde oder nicht, wurde dabei schon von der korrigierenden Person vorgenommen. Die Tatsache ob ein Punkt gegeben wurde oder nicht, bildet also das Selektionskriterium für die induktive Kategorienbildung. Außerdem muss bestimmt werden auf welchen Aspekt in weiterer Folge das Augenmerk gelegt

wird. Ist das Selektionskriterium einmal erfüllt liegt der Beobachtungsschwerpunkt auf Fehlern – diese sollen beschrieben werden. Auch dieser Punkt ist von nicht zu unterschätzender Wichtigkeit, wird nur das Selektionskriterium als Bezugspunkt herangezogen, könnte im Anschluss beispielsweise genau so gut die Handschrift oder die Farbe des verwendeten Stiftes kategorisiert werden. In weiterer Folge muss laut MAYRING (2015: 87) das Abstraktionsniveau der Kategorien bestimmt werden. Beispielsweise hat es keinen Sinn eine Kategorie *Konstruktionsfehler* zu bilden und weiters die Kategorie *Fehler bei der Winkelmessung* festzulegen. Den Prozess der induktiven Kategorienbildung, der auch in Abbildung 6 (MAYRING 2015: 84) dargestellt wird, beschreibt Philipp MAYRING (2015: 87) wie folgt:

Wenn das erste Mal das Selektionskriterium im Material erfüllt ist, wird möglichst nahe an der Textformulierung unter Beachtung des Abstraktionsniveaus die erste Kategorie als Begriff oder Kurzsatz formuliert. Wenn das nächste Mal das Selektionskriterium erfüllt ist, wird entschieden, ob die Textstelle unter die bereits gebildete Kategorie fällt (Subsumption) oder eine neue Kategorie zu bilden ist.

Wenn auf diese Weise ein großer Teil des Materials durchgearbeitet wurde und nur wenig neue Kategorien gebildet werden müssen (bei großen Materialmengen kann das schon bei zehn Prozent des Textes sein), ist dies der Moment für eine Revision des Kategoriensystems. Es muss überprüft werden, ob die Kategorien dem Ziel der Analyse nahekommen, ob das Selektionskriterium und das Abstraktionsniveau vernünftig gewählt worden sind. Wenn sich hier Veränderungen ergeben, muss mit der Analyse des Materials nochmals begonnen werden; ansonsten wird weitergegangen (nun kommen nur noch neue Kategorien dazu).

REINHOFFER (2008: 126) ergänzt: „Was nicht ins Kategorienschema passt, darf nach einer strengen Prüfung einer Rest-Kategorie zugeordnet werden.“

Nun ist immer die Rede von Text, nicht unbedingt der Ausdruck, der sofort mit der Bearbeitung von mathematischen Aufgaben in Verbindung gebracht wird. Aber bei Text handelt es sich immer nur um sprachliches Material und genau so wird Mathematik auch beschrieben. Bei dieser handelt es sich um eine schriftliche Sprache, die laut KLIX (1995: 33-34) zwei Funktionen haben kann: Zum einen die kommunikative Funktion und zum anderen die kognitive Funktion, wobei beide eng miteinander verbunden sind. Während, nach MAIER & SCHWEIGER (1999: 17-18), die kognitive Funktion dem Erkenntnisgewinn dient, trägt die kommunikative Funktion die kognitive Leistung nach außen. Im Falle der vorliegenden Daten erfüllt die Niederschrift (der Text) die kommunikative Funktion mathematischer Sprache, welche wiederum Auskunft über die kognitive Leistung gibt.

Am Ende der induktiven Kategorienbildung steht ein System an Kategorien, die oft mit konkreten Ankerbeispielen verbunden sind. Die fertig ausformulierten Kategorien können sowohl Ende als auch Anfang sein. MAYRING (2015: 87) nennt drei Möglichkeiten wie mit dem entstandenen Kategoriensystem weiter gearbeitet werden kann:

- Das ganze Kategoriensystem kann interpretiert werden im Sinne der Fragestellung.
- Es können entweder induktiv (im Sinne zusammenfassender Inhaltsanalyse) oder deduktiv (mithilfe theoretischer Erwägungen) Hauptkategorien gebildet werden.
- Es können quantitative Analysen, z.B. Häufigkeiten der Kategorien angefügt werden.

Schlussendlich stehen auch im Rahmen von dieser Arbeit Hauptkategorien (*Kapitel 4*), die wiederum induktiv gebildet werden. Diese Hauptkategorien werden das Abstraktionsniveau eine Stufe hinaufsetzen. Gilt es zunächst Kategorien auf der Ebene der Aufgaben zu finden (*Abschnitt 3.3*), sollen die Hauptkategorien aufgabenübergreifend entstehen. Vorher werden die entstandenen Kategorien quantitativ analysiert. Wie MAYRING vorschlägt, werden dabei die Häufigkeiten der Ausprägungen der jeweiligen Kategorien eruiert und festgehalten.

3.2 Häufigkeitsanalyse

Nachdem die deskriptiven Kategorien gefunden wurden, stellt sich die Frage, was mit ihnen passieren soll. Die Antwort ist denkbar einfach: Es wird ermittelt wie oft etwas unter einer Kategorie subsumiert werden konnte – es wird also eine Häufigkeitsanalyse durchgeführt. MAYRING (2015: 13) bezeichnet den Prozess „bestimmte Elemente des Materials auszuzählen und in ihrer Häufigkeit mit dem Auftreten anderer Elemente zu vergleichen“ als „die einfachste Art inhaltsanalytischen Arbeitens“. Dabei wird zunächst ein Kategoriensystem gebildet (im vorliegenden Fall durch induktive Kategorienbildung), woraufhin die Häufigkeiten der verschiedenen Kategorien festgestellt und verglichen werden können (MAYRING 2015: 15).

Bevor mit den Häufigkeiten weitergearbeitet werden kann, muss zuerst festgehalten werden, um welche Daten es sich handelt. HOLLAND & SCHARNBACHER (2010: 5-6) unterscheiden zwischen vier Skalen: Die *Nominalskala*, die *Ordinalskala*, die *Intervallskala* und die *Verhältnisskala*. Bei der Nominalskala kann zwar zwischen den Merkmalsausprägungen unterschieden werden, allerdings können diese nicht in eine Reihenfolge gebracht werden. Die Ordinalskala unterscheidet sich von der Nominalskala dadurch, dass ordinal skalierte Daten geordnet werden können. Eine Aussage über die Abstände der geordneten Daten lässt sich aber nicht treffen. Dieser Punkt folgt erst bei der Intervallskala. Während nun bei der Intervallskala der Nullpunkt willkürlich festgelegt wurde, liegt der Verhältnisskala ein absoluter Nullpunkt zugrunde. Die Intervall- und die Verhältnisskala werden auch unter dem Begriff der Kardinalskala zusammengefasst (TIEMANN 2012: 22).

Die deskriptiven Kategorien, die aus den vorliegenden Daten entstehen, können nicht sinnvoll geordnet werden, womit auch eine Abstandsmessung nicht möglich ist. Bei den Kategorien beziehungsweise den Merkmalsausprägungen handelt es sich also um nominal skalierte Daten. Die Häufigkeiten einer Nominalskala lassen sich sowohl in *Tabellen* als auch mittels *Stabdiagrammen* darstellen. Bei einer Tabelle sind laut TIEMANN (2012: 25) zumindest die Merkmalsausprägungen, die absoluten Häufigkeiten sowie die relativen Häufigkeiten darzustellen. Je nach Art der Daten sind auch die kumulierten relativen Häufigkeiten zu nennen. Mit der Darstellung dieser Merkmale sieht TIEMANN (2012: 25-26, Hervorhebung im Original) die Aufgabe von Häufigkeitstabellen als erfüllt an: „Die Häufigkeitstabelle sammelt alle verschiedenen Merkmalsausprägungen, zählt, wie oft jede vorkommt, und stellt das

Ergebnis tabellarisch dar. [...] Mit Hilfe der Häufigkeitstabelle können *Anteilsfragen* gestellt werden – und mit Hilfe der passenden Spalte schnell beantwortet werden“.

Mithilfe des Stabdiagramms können Häufigkeitstabellen graphisch dargestellt werden. Meist wird dabei die horizontale Achse als Merkmalsachse verwendet, während auf der vertikalen Achse die (relativen) Häufigkeiten proportional abgetragen werden, so TIEMANN (2012: 27). Nach demselben Prinzip werden *Säulendiagramme* sowie *Balkendiagramme* erstellt. Im Unterschied zum Stabdiagramm sind die senkrecht abgetragenen Strecken beim Säulendiagramm etwas breiter, beim Balkendiagramm handelt es sich um ein um 90° im Uhrzeigersinn gedrehtes Balkendiagramm. Aus all diesen Diagrammen lässt sich ein Wert besonders gut herauslesen: Der sogenannte *Modus*, der die häufigste Merkmalsausprägung bezeichnet.

Sowohl für die Häufigkeitstabelle als auch für jedwede graphische Darstellung gilt, was für jede Zusammenfassung gilt: Kein Überblicksgewinn ohne Informationsverlust. Auch TIEMANN (2012: 30) betont nochmals, dass „Häufigkeiten [...] bei nahezu jeder Datenlage [gebildet werden können], ggf. gehen Ihnen dabei allerdings wichtige Erkenntnisse verloren“.

3.3 Exemplarische Auswertung

Bevor es an die Auswertung der gewählten zwölf Aufgaben geht (siehe *Abschnitt 2.1.2*), wird eine Aufgabe exemplarisch bearbeitet, um die Vorgehensweise der induktiven Kategorienbildung und der Häufigkeitsanalyse detailliert zu zeigen. Dazu werden die einzelnen Schritte des Prozessmodells induktiver Kategorienbildung (Abbildung 6) durchlaufen, woraufhin die gebildeten Kategorien auf ihre Häufigkeiten untersucht werden.

Die Grundlage für jede Untersuchung ist das zugrundeliegende Material, wie bereits erwähnt, handelt es sich hier um die bearbeiteten Aufgaben von 301 Gymnasialschülerinnen und -schülern der ssRP Mathematik 2016 (siehe *Abschnitt 2.1.2*). Das Ziel ist es, die Daten nutzbar zu machen, also die Antworten zu kategorisieren. Genauer gesagt, sollen die Fehler kategorisiert werden, das Kategoriensystem kommt also nur zum Einsatz, wenn der Punkt nicht gegeben, die Aufgabe also falsch bearbeitet wurde. Dies stellt zugleich das Selektionskriterium dar. Tritt dieses in Kraft wird die jeweilige Antwort analysiert und einer Kategorie zugeordnet oder eine neue Kategorie gebildet. Die richtigen Antworten bilden ebenfalls eine eigene Kategorie.

Bei der nun im Detail untersuchten Aufgabe handelt es sich um die Aufgabe 5 (kurz: X5) des Haupttermins 2016 (BIFIE 2016a: 9), eine Aufgabe aus dem Gebiet *Algebra und Geometrie* (siehe auch Abbildung 7). Sie setzt sich mit der Grundkompetenz AG 3.3 auseinander: „Definition der Rechenoperationen mit Vektoren (Addition, Multiplikation mit einem Skalar, Skalarmultiplikation) kennen, Rechenoperationen verständlich einsetzen können und (auch geometrisch) deuten können“ (BIFIE 2013a: 7).

Vektoraddition

Die unten stehende Abbildung zeigt zwei Vektoren $\vec{v}_1$ und $\vec{v}$.

Aufgabenstellung:

Ergänzen Sie in der Abbildung einen Vektor $\vec{v}_2$ so, dass $\vec{v}_1 + \vec{v}_2 = \vec{v}$ ist!

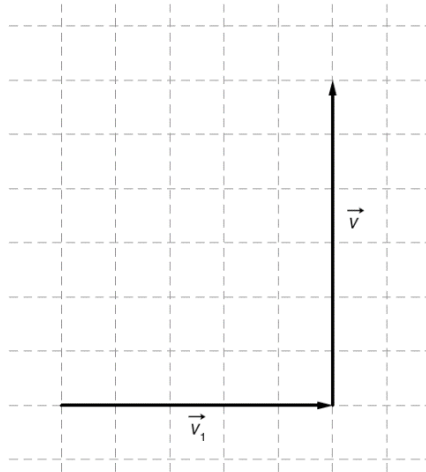


Abbildung 7: Aufgabe 5 (X5)

Mit der Aufgabenstellung beginnt die Kategorisierung. Die vorliegende Aufgabe ist dem Konstruktionsformat zuzuordnen. Nachdem die Kategorien nach MAYRING (2015: 86) möglichst „in der Sprache des Materials“ gehalten werden sollen, hat es Sinn die Kategorien ebenfalls mit Konstruktionen darzustellen. Mithilfe der Definitionstheorie können die Kategorien zwar exakt festgelegt werden, jedoch müssten dann auch alle Entsprechungen der Kategorien gleich gute Beispiele für eben jene Kategorie sein. Da dies bei 301 verschiedenen Konstruktionen kaum der Fall sein kann, hat es mehr Sinn die Kategorien über Prototypen zu bestimmen.

Mit der ersten Performanz wird folgerichtig auch die *erste Kategorie (i)* gebildet⁶ (Abbildung 8). In diesem Falle beinhaltet diese Kategorie alle Antworten, die einen Vektor $\vec{v}_2$ beschreiben, für den gilt $\vec{v}_2 = \vec{v} - \vec{v}_1$. Ob diese Kategorie Sinn ergibt, ob noch weitere

⁶ Die Abbildungen zu den Kategorien wurden entsprechend den Performanzen der Schülerinnen und Schülern nachgezeichnet und so anonymisiert.

Entsprechungen folgen, lässt sich zum jetzigen Zeitpunkt noch nicht feststellen. Dies kann erst im nächsten Schritt des Prozessmodells, der Revision, ermittelt werden.

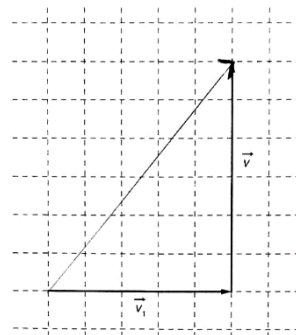


Abbildung 8: X5 Kategorie i

Danach wird Aufgabenbearbeitung für Aufgabenbearbeitung fortgesetzt. Einer weiteren Antwort, die sich Kategorie i zuordnen lässt, folgen zwei korrekt gelöste Performanzen. Die darauffolgende Antwort lässt sich nicht in die erste Kategorie subsumieren, folglich muss eine neue Kategorie gebildet werden. Algebraisch gesprochen entspricht die *Kategorie ii* (Abbildung 9) dem Vektor $\vec{v}_2 = -\vec{v}_1 - \vec{v}$.

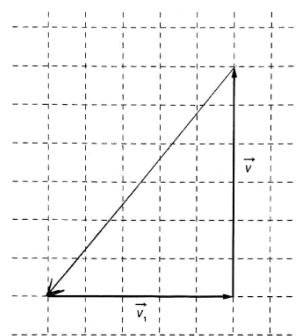


Abbildung 9: X5 Kategorie ii

Mit diesen beiden Kategorien kann eine ganze Weile gearbeitet werden. Die Antworten sind entweder Kategorie i oder Kategorie ii zuzuordnen, oder wurden richtig gelöst. Performanzen, die keine Antwort enthalten, können auch keiner (bisher bestehenden) Kategorie zugeordnet werden und so wird die selbsterklärende *Kategorie Nicht gemacht (NG)* gebildet. Der Prozess der Materialdurcharbeitung läuft fort und fördert zwei neue Kategorien zu Tage. Zum einen die *Kategorie iii* (Abbildung 10), die jene Antworten beschreibt, die zwar den korrekten Richtungsvektor aufweisen, der Betrag des Vektors allerdings zu groß ist. Zum anderen die *Kategorie iv* (Abbildung 11), bei der keine Orientierung erkennbar ist, der Richtungsvektor allerdings entweder derselbe wie in Kategorie i oder Kategorie ii sein muss.

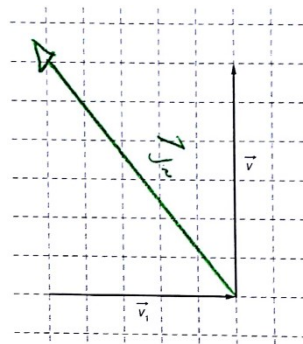


Abbildung 10: X5 Kategorie iii

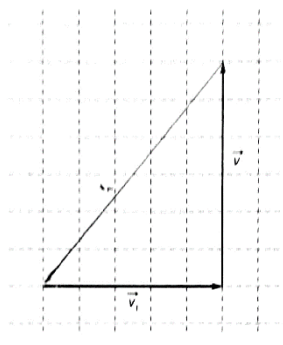


Abbildung 11: X5 Kategorie iv

Keiner bisher beschriebenen Kategorie kann eine Antwort der algebraischen Form $\vec{v}_2 = \vec{v}$ zugeordnet werden. Diese bildet folglich die nächste *Kategorie v* (Abbildung 12). Bevor es an die Revision geht, muss noch eine letzte Kategorie gebildet werden, die mit $\vec{v}_2 = -\frac{2}{5}\vec{v}_1 - \vec{v}$ beschrieben werden kann und die finale *Kategorie vi* (Abbildung 13) bildet.

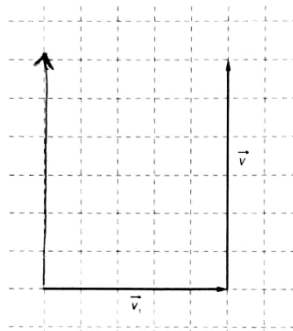


Abbildung 12: X5 Kategorie v

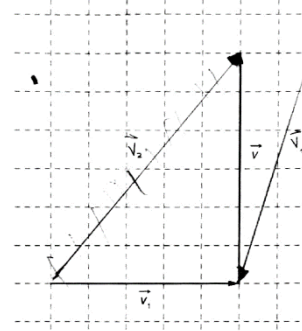


Abbildung 13: X5 Kategorie vi

Bis zu dieser Kategorie vi wurden 100 Performanzen ausgewertet, also rund ein Drittel der vorliegenden Arbeiten. Ein guter Zeitpunkt kurz innezuhalten und die bisher entstanden Kategorien nochmals zu betrachten und zu bewerten. Zum momentanen Zeitpunkt werden sechs verschiedene Kategorien geführt (Abbildung 14).

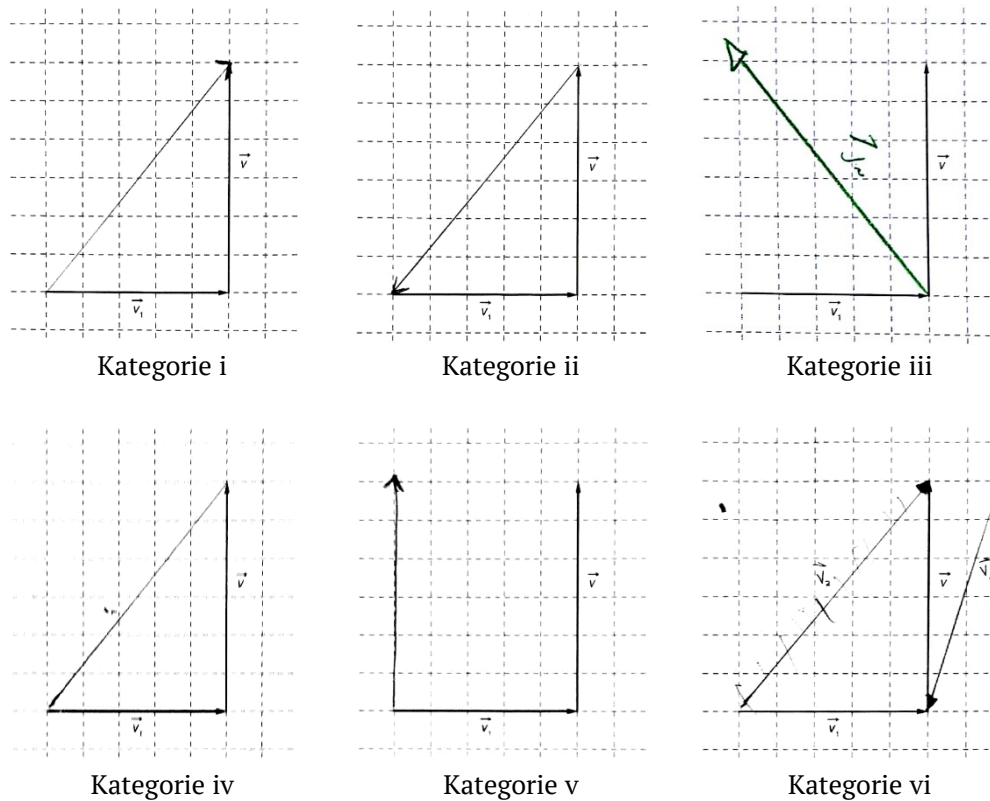


Abbildung 14: X5 Vorläufiges Kategoriensystem

Bei der ersten Kategorie handelt es sich um die bislang häufigste mit 26 Entsprechungen. Für Kategorie ii gibt es zwölf Entsprechungen, für Kategorie iii eine, für Kategorie iv ebenfalls eine, für Kategorie v vier und für Kategorie vi gibt es wieder eine Entsprechung. Es existieren somit nur drei Kategorien, die mehr als drei Entsprechungen aufweisen können. Die anderen drei Kategorien werden nun in einer *Restkategorie (RK)* zusammengefasst. Zusätzlich kommt noch die *Kategorie Nicht gemacht (NG)* hinzu, die allerdings aus nachvollziehbaren Gründen keine Veranschaulichung benötigt. Somit ergibt sich das in Abbildung 15 dargestellte Kategoriensystem:

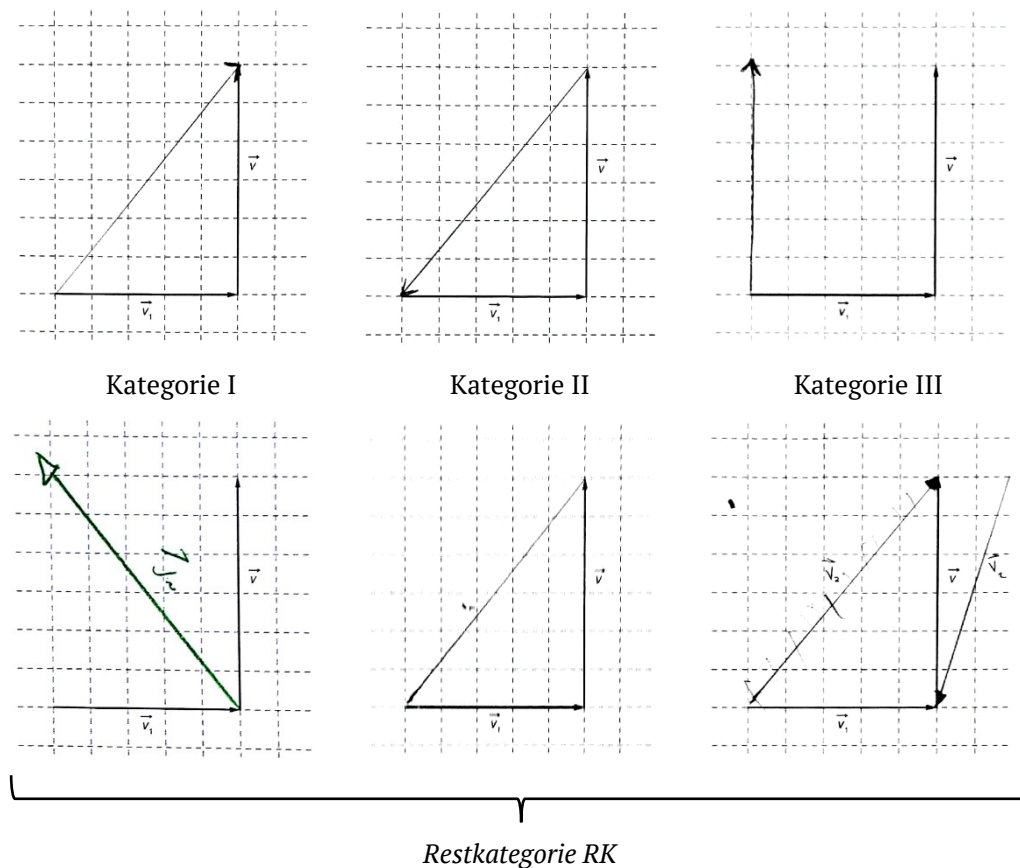


Abbildung 15: X5 Kategoriensystem

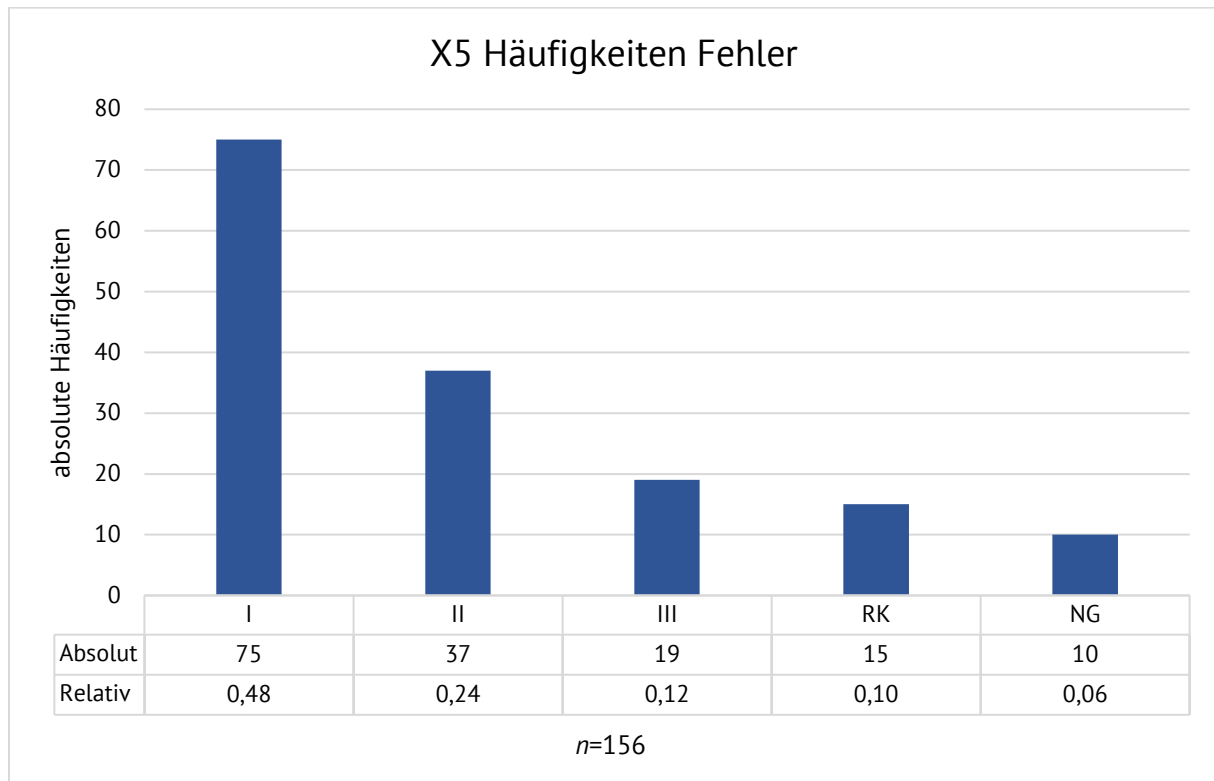
Von nun an wird mit dem in Abbildung 15 dargestellten Kategoriensystem, zuzüglich der *Kategorie Nicht Gemacht*, weiter gearbeitet. Tauchen weitere Antworten auf, die sich nur der *Restkategorie* zuordnen lassen, muss geprüft werden, ob dadurch nicht eine Häufung einer neuen Kategorie innerhalb der *Restkategorie* entsteht. Ist dies der Fall muss eine neue Kategorie gebildet werden, ist dies nicht der Fall verbleiben diese Antworten in der *Restkategorie*.

Nach Durchsicht aller Performanzen hat sich das Kategoriensystem bestätigt. Keine neuen Kategorien mussten hinzugefügt werden. Die *Restkategorie* musste zwar noch einige Male angewandt werden (insgesamt 15 Mal), allerdings hätten jeweils nur maximal zwei Performanzen zu einer Kategorie zusammengefasst werden können. Das nach der Revision beschriebene Kategoriensystem bleibt also bestehen und umfasst die Kategorien I, II, III, RK und NG. Insgesamt wurde das Selektionskriterium 156 Mal ausgelöst, d. h. die 156 falschen Antworten wurden kategorisiert.

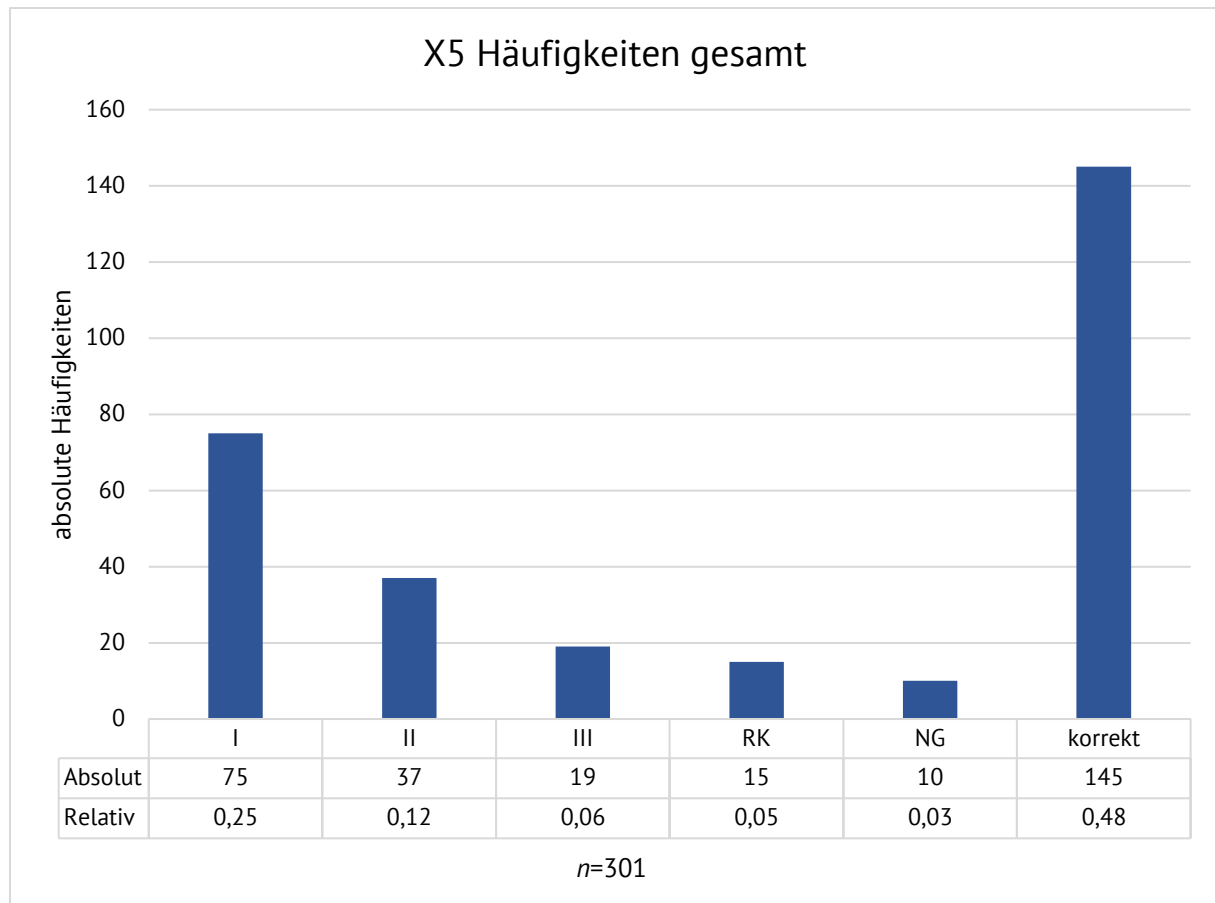
Im nächsten Schritt werden die Kategorien quantitativ untersucht. Dazu werden die Entsprechungen zu den einzelnen Kategorien gezählt. Dies kann sowohl in einer Häufigkeitstabelle als auch mit Hilfe eines Säulendiagramms dargestellt werden. Zunächst

werden die Kategorien in Bezug auf die fehlerhaften Antworten gesetzt. Statistik 1 stellt folglich den Anteil einer bestimmten Kategorie an allen fehlerhaften Antworten (hier: $n=156$) dar. In der darauffolgenden Darstellung (Statistik 2) werden die Anteile der Kategorien an allen Performanzen ($n=301$, inklusive der korrekten Antworten) dargestellt. Die korrekt gelösten Antworten bilden demnach eine weitere Kategorie.

Auswertung



Statistik 1: X5 Grundgesamtheit: fehlerhafte Antworten



Statistik 2: X5 Grundgesamtheit: alle Performanzen

Aus Statistik 1 lässt sich leicht der Modus erkennen, die Kategorie I. In ungefähr 48 Prozent aller falsch gelösten Aufgaben wurde dieselbe Strategie verfolgt. Mithilfe von Statistik 2 kann der Anteil einer Kategorie an allen Performanzen ermittelt werden. Im Falle der Kategorie I wird klar, dass ein Viertel aller Schülerinnen und Schüler die Aufgabe entsprechend bearbeitet haben. Die zweithäufigste Kategorie (Kategorie II) enthält nur halb so viele Entsprechungen wie Kategorie I, 37 Schülerinnen und Schüler haben diesen Lösungsansatz verfolgt. Auch die Häufigkeiten der übrigen Kategorien lassen sich leicht aus den beiden Darstellungen – jeweils für die Grundgesamtheit der fehlerhaften Antworten und die Grundgesamtheit aller Performanzen – herauslesen. Insgesamt haben somit 145 Masantinnen und Masanten (etwa 48 Prozent) die Aufgabe richtig gelöst, während 156 Schülerinnen und Schüler (etwa 52 Prozent) die Aufgabenstellung falsch bearbeitet haben.

Eine abschließende Bemerkung gilt es noch zu machen. Während in Abbildung 2 die Lösungsquote der fünften Aufgabe mit 49 Prozent angegeben wurde, ist hier nur mehr von 48 Prozent richtig gelösten Performanzen die Rede. Für die Erstellung des Diagramms der Lösungsquoten in Abbildung 2 (S. 11) wurden die offiziellen Daten des BMB herangezogen. Nach der Untersuchung der Aufgaben im Detail stellte sich heraus, dass vereinzelt Punkte

vergeben wurden, die laut Lösungserwartung (BIFIE 2016b) nicht zu vergeben gewesen wären und somit einer der Fehlerkategorien zuzuordnen waren. Das Diagramm zu den durchschnittlichen Lösungsquoten auf Seite 11 basiert also auf den tatsächlich vergebenen Punkten, während die Häufigkeitstabellen und Diagramme zu den einzelnen Aufgaben die Häufigkeiten der tatsächlich korrekt bearbeiteten Aufgaben dokumentieren. Dies erklärt die Diskrepanz zwischen den beiden Werten. Wenn nicht anders angegeben, lässt sich ein möglicher Unterschied der durchschnittlichen Lösungsquoten zwischen Abbildung 2 und den jeweiligen Häufigkeitstabellen auf diese Weise begründen.

4 Auswertung

In diesem Kapitel geht es an die konkrete Auswertung der gewählten Aufgaben. Wie in *Abschnitt 2.1* bereits beschrieben, handelt es sich dabei um die Items 1-6 sowie 19-24 und somit um die Typ-1-Aufgaben zu den Themen *Algebra und Geometrie* sowie *Wahrscheinlichkeit und Statistik*. Diese Aufgaben werden nach den in *Kapitel 3* vorgestellten Arbeitsweisen bearbeitet: In einem ersten Schritt werden Fehlerkategorien gebildet, deren Häufigkeiten in weiterer Folge mithilfe von Häufigkeitstabellen und / oder Säulendiagrammen dargestellt werden. Bei den angegebenen Werten handelt es sich stets um gerundete Angaben. Das Design der Auswertung gestaltet sich wie folgt:

Zunächst wird zu jedem Item die Aufgabenstellung zitiert. Diese entstammt dem Haupttermin aus Mathematik 2016 des BIFIE (2016a). Daraufhin wird die, mit der Aufgabe korrespondierende, Grundkompetenz genannt. Die Grundkompetenzen werden ebenfalls nach Definitionen des BIFIE (2013a) wiedergegeben und wurden vom BMB den jeweiligen Aufgaben zugeordnet. Nach diesen beiden Angaben folgt eine kurze Beschreibung der Aufgabe, bevor die Kategorien erklärt werden. Diese werden in einer Häufigkeitstabelle beziehungsweise einem Säulendiagramm zusammengefasst – die Häufigkeiten der Entsprechungen werden erfasst. Des Weiteren werden etwaige Bemerkungen oder Besonderheiten zur jeweiligen Aufgabe genannt.

Exkurs: Antwortformate

Vor der ersten zentral gestellten Mathematikmatura im Jahr 2014 weist das BIFIE (2013a: 26) darauf hin, dass „einige Bereiche [...] anders als es die Schüler/innen bei der Bearbeitung von Aufgaben im Unterricht, z. B. im Rahmen von Schularbeiten, gewohnt sind“ seien. Von ursprünglich acht verschiedenen (Aufgaben-)Formaten, von denen 2013 die Rede ist (BIFIE 2013a: 26), treten bei der ssRP Mathematik 2016 sieben verschiedene Formate auf. Dabei handelt es sich um *freie Antwortformate*, das *Zuordnungsformat*, das *Konstruktionsformat*, das *Multiple-Choice-Format in der Variante 1 aus 6*, das *Multiple-Choice-Format in der Variante 2 aus 5*, das *Multiple-Choice-Format in der Variante x aus 5* und den *Lückentext*. In der Erläuterung der Antwortformate zur ssRP (BIFIE 2016a: 2-4) wird nicht mehr zwischen offenen und halboffenen Formaten unterschieden. Die Aufgaben können dennoch den ursprünglichen acht Formaten zugeordnet werden, um eine bessere Beschreibung gewährleisten zu können.

Die Verteilung der verschiedenen Antwortformate auf die untersuchten Aufgaben ist in Tabelle 5 zu sehen.

Format	Aufgabe
2 aus 5	X1
Offen	X2, X4, X6, X20, X21
Halboffen	X3, X22
Konstruktion	X5, X19
1 aus 6	X23
Lückentext	X24

Tabelle 5: Antwortformate untersuchte Aufgaben

Diese sechs Antwortformate werden vom BIFIE (2013a: 27-31) wie folgt beschrieben:

1. Multiple-Choice-Aufgabenformat: 2 aus 5
Dieses Antwortformat ist durch einen Fragenstamm und fünf Antwortmöglichkeiten gekennzeichnet, wobei zwei Antwortmöglichkeiten auszuwählen sind. Aufgaben dieses Formats werden korrekt bearbeitet, indem die beiden zutreffenden Antwortmöglichkeiten angekreuzt werden.
2. Offenes Antwortformat
Die Antwort soll mit eigenen Worten formuliert werden bzw. darf völlig frei erfolgen.
3. Halboffenes Antwortformat
Die korrekte Antwort oder ein vorgegebenes bzw. passendes mathematisches Objekt soll in eine vorgegebene Formel, Funktion etc. eingesetzt werden.
4. Konstruktionsformat
Hier sollen in ein vorgegebenes Koordinatensystem (dessen Achsenskalierung nicht standardisiert ist) entsprechende Graphen, Punkte, Vektoren o. Ä. eingetragen werden.
Eine Aufgabe und deren Aufgabenstellung sind vorgegeben. Die Aufgabe erfordert die Ergänzung von Punkten, Geraden und/oder Kurven im Aufgabenheft.
5. Multiple-Choice-Aufgabenformat: 1 aus 6
Dieses Antwortformat ist durch einen Fragenstamm und sechs Antwortmöglichkeiten gekennzeichnet, wobei eine Antwortmöglichkeit auszuwählen ist. Aufgaben dieses Formats werden korrekt bearbeitet, indem die zutreffende Antwortmöglichkeit angekreuzt wird.
6. Lückentext
Dieses Antwortformat ist durch einen Satz mit zwei Lücken gekennzeichnet, das heißt, im Aufgabentext sind zwei Stellen ausgewiesen, die ergänzt werden müssen. Für jede Lücke werden je drei Antwortmöglichkeiten vorgegeben. Aufgaben dieses Formats werden korrekt bearbeitet, indem die Lücken durch Ankreuzen der beiden zutreffenden Antwortmöglichkeiten gefüllt werden.

Exkurs Ende

4.1 Aufgabe 1

Menge von Zahlen

Die Menge $M = \{x \in \mathbb{Q} \mid 2 < x < 5\}$ ist eine Teilmenge der rationalen Zahlen.

Aufgabenstellung:

Kreuzen Sie die beiden zutreffenden Aussagen an!

4,99 ist die größte Zahl, die zur Menge M gehört.	☐
Es gibt unendlich viele Zahlen in der Menge M , die kleiner als 2,1 sind.	☐
Jede reelle Zahl, die größer als 2 und kleiner als 5 ist, ist in der Menge M enthalten.	☐
Alle Elemente der Menge M können in der Form $\frac{a}{b}$ geschrieben werden, wobei a und b ganze Zahlen sind und $b \neq 0$ ist.	☐
Die Menge M enthält keine Zahlen aus der Menge der komplexen Zahlen.	☐

1. Grundkompetenz: AG 1.1

Wissen über die Zahlenmengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ verständig einsetzen können

2. Beschreibung

Bei der ersten Aufgabe (X1) des Mathematik Haupttermins 2016 handelt es sich um eine Aufgabe des Formats *Multiple-Choice-Format in der Variante 2 aus 5*. Die beiden richtigen Antworten sind II und IV, wobei genau diese beiden angekreuzt werden müssen, damit die Aufgabe als richtig beantwortet gewertet wird. Werden die Kategorien, die ganz „in der Sprache des Materials“ (MAYRING 2015: 86) bleiben, zusammengefasst, ergibt sich Tabelle 6.

3. Kategorien

Kategorie	Beschreibung
I	4,99 ist die größte Zahl, die zur Menge M gehört.
II	Es gibt unendlich viele Zahlen in der Menge M , die kleiner als 2,1 sind.
III	Jede reelle Zahl, die größer als 2 und kleiner als 5 ist, ist in der Menge M enthalten.
IV	Alle Elemente der Menge M können in der Form $\frac{a}{b}$ geschrieben werden, wobei a und b ganze Zahlen sind und $b \neq 0$ ist.
V	Die Menge M enthält keine Zahlen aus der Menge der komplexen Zahlen.

Tabelle 6: X1 Kategoriensystem

Infolge des Formats lassen sich die Antworten der Schülerinnen und Schüler recht leicht kategorisieren: Bei den Kategorien handelt es sich um die anzukreuzenden Aussagen. Diese werden dabei, mit der ersten beginnend, mit römischen Zahlen bezeichnet (Tabelle 6). So entspricht Abkürzung III beispielsweise der dritten Antwort. Die Kategorien können somit mit folgender Zuordnungsvorschrift klar voneinander abgetrennt werden: Eine Antwort ist genau dann einer *Kategorie* (*I*, *II*, *III*, *IV*, *V*) zuzuordnen, wenn das entsprechende Kreuz gesetzt wurde. Folgerichtig wurde hier die Definitionstheorie angewandt (siehe *Abschnitt 3.1*), die idealerweise, und hier tatsächlich, in diskreten Kategorien mündet. Mit dem besonderen Fall der Multiple-Choice-Aufgabe werden also auch WALDMANNs (2016: 381) Bedenken entkräftet, dass Kategorien in den seltensten Fällen wirklich diskret sind.

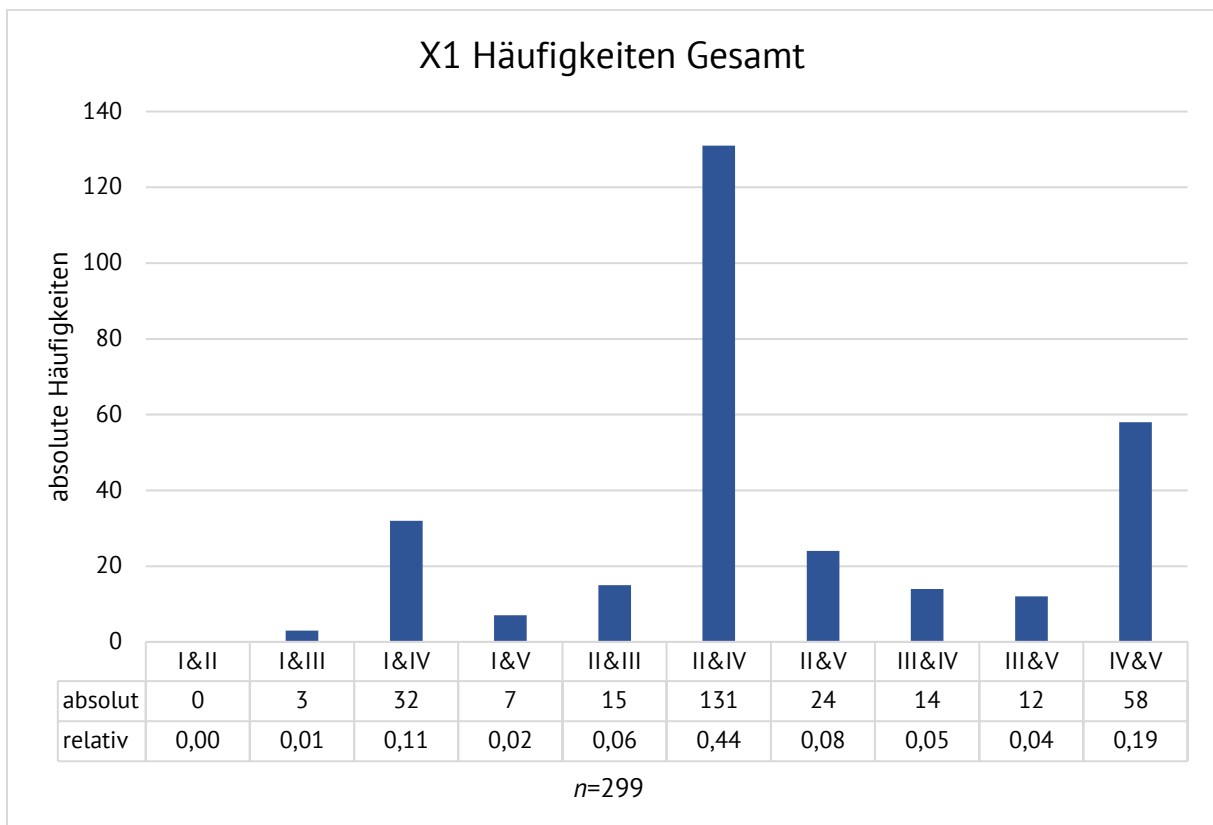
4. Auswertung

Absolut (n=299)					
I	II	III	IV	V	&
-	0	3	32	7	I
	-	18	131	24	II
		-	14	12	III
			-	58	IV
				-	V

Tabelle 7: X1 Häufigkeitstabelle (absolut) verknüpfte Antworten (Grundgesamtheit: alle Performanzen)

Relativ (n=299)					
I	II	III	IV	V	&
-	0,00	0,01	0,11	0,02	I
	-	0,06	0,44	0,08	II
		-	0,05	0,04	III
			-	0,19	IV
				-	V

Tabelle 8: X1 Häufigkeitstabelle (relativ) verknüpfte Antworten (Grundgesamtheit: alle Performanzen)



Statistik 3: X1 Grundgesamtheit: alle Performanzen

n=299	I	II	III	IV	V
Absolut	42	173	47	235	101
Relativ	0,14	0,58	0,16	0,79	0,34

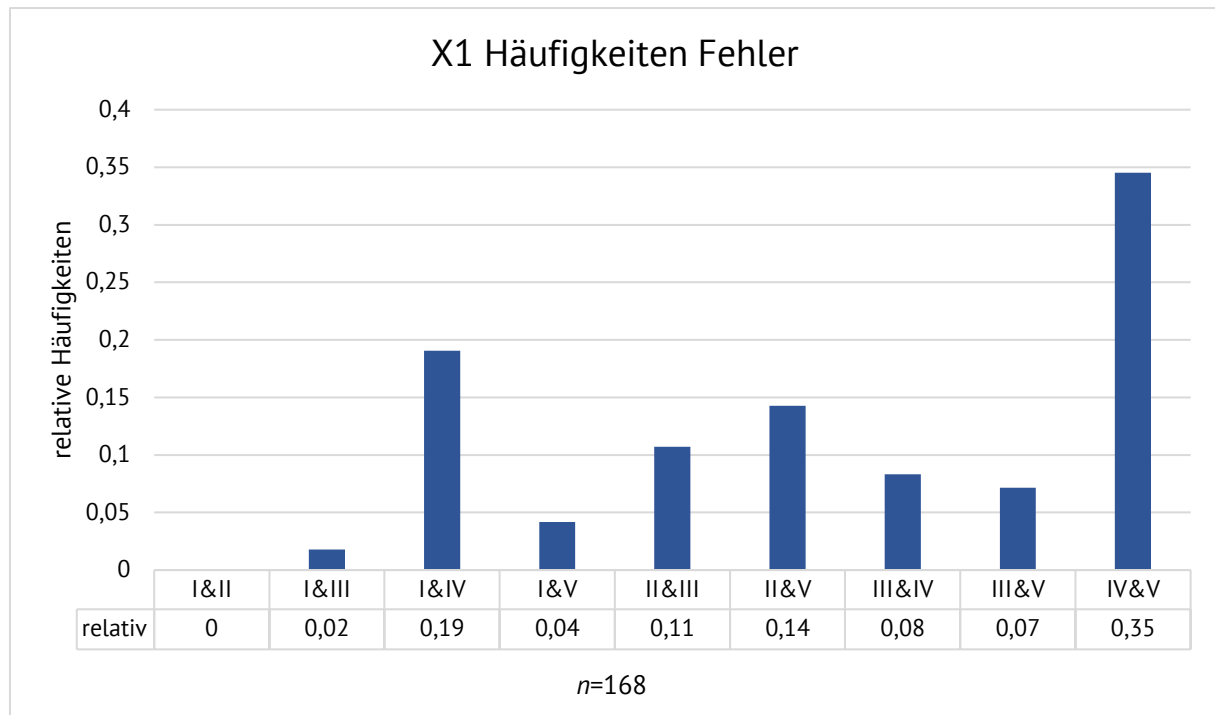
Tabelle 9: X1 Häufigkeitstabelle einzelne Ankreuzungen (Grundgesamtheit: alle Performanzen)

Gleich bei der Auswertung der ersten Aufgabe ergibt sich eine Besonderheit: Bisher war immer von einer Grundgesamtheit von 301 Schülerinnen und Schülern die Rede, nun wird als Grundgesamtheit $n=299$ angegeben. Zwei Schülerinnen oder Schüler haben drei Kästchen angekreuzt, woraufhin diese beiden Performanzen bei den verknüpften Tabellen (Tabelle 7 und Tabelle 8) nicht berücksichtigt werden konnten. In Tabelle 9, die die Anzahl der einzelnen Ankreuzungen beschreibt (Mehrfachnennungen möglich), hätten zwar auch diese beiden Performanzen wieder berücksichtigt werden können, allerdings wäre dann die Vergleichbarkeit zwischen Tabelle 7 und Tabelle 9 nicht mehr gegeben, weshalb weiterhin mit derselben Grundgesamtheit gearbeitet wird.

Aus Tabelle 9 wird ersichtlich, dass die vierte Antwort mit Abstand am häufigsten angekreuzt wurde. 235 Schülerinnen und Schüler (79 Prozent) haben somit zumindest eine richtige Antwort erkannt. Vergleicht man diese Anzahl mit der Anzahl der tatsächlich korrekt gelösten Aufgaben (131) aus Tabelle 7 ergibt sich eine große Differenz: 104 Schülerinnen und Schüler ($235 - 131 = 104$) und somit fast 35 Prozent haben zwar eine korrekte Aussage erkannt, jedoch nicht die zweite. Nun stellt sich die Frage mit welchen Kombinationen die falschen Antworten einhergehen. Um dies festzustellen, wird die Grundgesamtheit der falschen Antworten ($n=168$) herangezogen und wiederum eine Tabelle der verknüpften Antworten erstellt (Tabelle 10). Die absoluten Zahlen bleiben dieselben wie in Tabelle 7.

Relativ ($n=168$)					
I	II	III	IV	V	&
-	0,00	0,02	0,19	0,04	I
	-	0,11	-	0,14	II
		-	0,08	0,07	III
			-	0,35	IV
				-	V

Tabelle 10: X1 Häufigkeitstabelle (relativ) verknüpfte Antworten (Grundgesamtheit: fehlerhafte Antworten)



Statistik 4: X1 Grundgesamtheit: fehlerhafte Antworten

Werden Tabelle 10 und Statistik 4 genauer betrachtet, wird klar, dass die häufigste Antwortkombination – aus den falsch bearbeiteten Performanzen – jene aus der (richtigen) vierten und der (falschen) fünften Antwort ist. 35 Prozent all jener Schülerinnen und Schüler, die die Aufgabe nicht korrekt gelöst hatten, haben denselben Fehler begangen. Die häufigste falsche Kombination ausgehend von einer Kategorie ist jene mit IV & eine beliebige Antwort (ausgenommen IV & II natürlich). Konkret heißt das, war eine Antwort falsch, wurde ein Kreuz in 62 Prozent der Fälle bei der vierten Antwort gemacht ($IV\&I + IV\&III + IV\&V = 0,19 + 0,08 + 0,35 = 0,62$). Dies lässt sich auch für die anderen Kategorien erstellen (Tabelle 11). Allerdings gilt es dabei zu beachten, dass Mehrfachnennungen vorkommen: bei der Häufigkeit für I&x kommt I&IV genauso vor wie bei IV&x. Des Weiteren sind die richtigen Kombinationen naturgemäß ausgeschlossen, sprich bei II&x und IV&x sind die jeweilig korrekten dazugehörigen Antworten (IV beziehungsweise II) nicht berücksichtigt.

$n=168$	I&x	II&x	III&x	IV&x	V&x
Relativ	0,25	0,25	0,28	0,62	0,60

Tabelle 11: X1 Häufigkeitstabelle Kategorien in falscher Kombination (Grundgesamtheit: fehlerhafte Antworten)

4.2 Aufgabe 2

Äquivalenzumformung

Nicht jede Umformung einer Gleichung ist eine Äquivalenzumformung.

Aufgabenstellung:

Erklären Sie konkret auf das unten angegebene Beispiel bezogen, warum es sich bei der durchgeführten Umformung um keine Äquivalenzumformung handelt! Die Grundmenge ist die Menge der reellen Zahlen.

$$\begin{array}{l} x^2 - 5x = 0 \quad | : x \\ x - 5 = 0 \end{array}$$

1. Grundkompetenz: AG 1.2

Wissen über algebraische Begriffe angemessen einsetzen können: Variable, Terme, Formeln, (Un-)Gleichungen, Gleichungssysteme, Äquivalenz, Umformungen, Lösbarkeit

Anmerkung: Die algebraischen Begriffe soll man anhand von einfachen Beispielen beschreiben/erklären und verständlich verwenden können.

2. Beschreibung

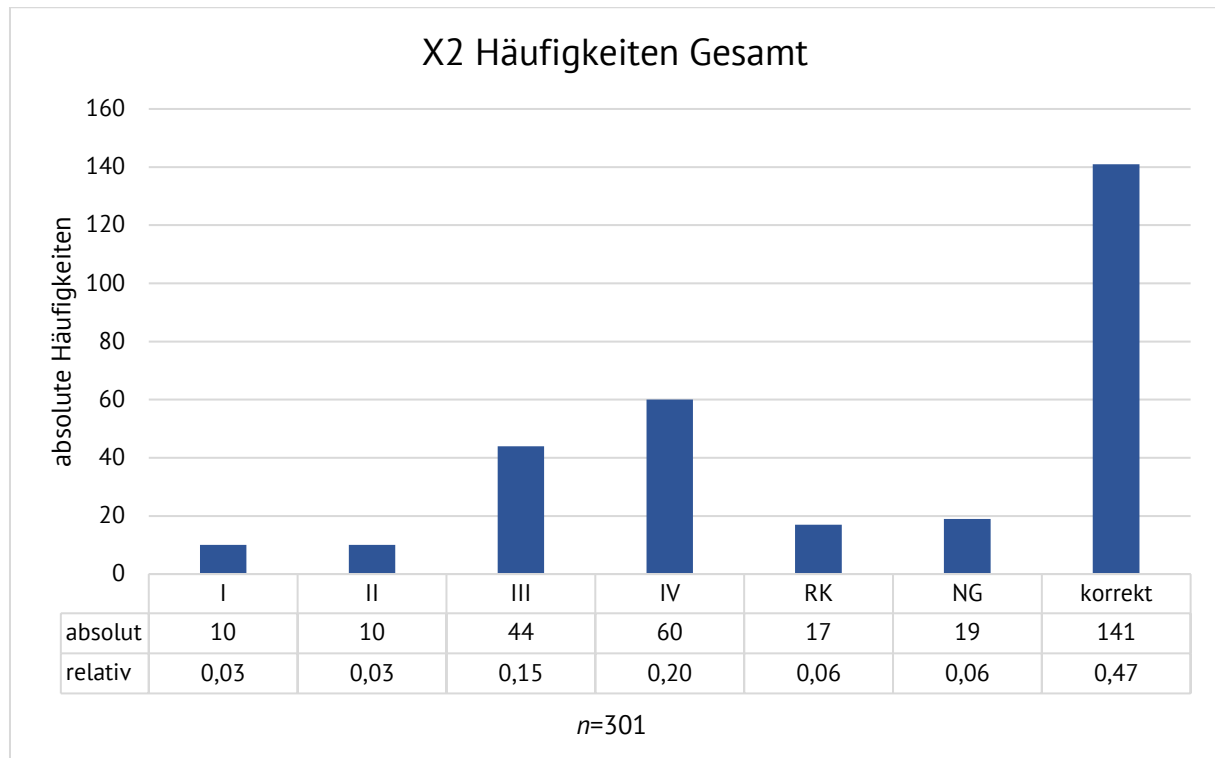
Bei Aufgabe 2 (X2) handelt es sich um die einzige der untersuchten Aufgaben, bei der die Lösungserwartung zwei mögliche Erklärungen angibt: Die erste Begründung beruht darauf, dass durch die vorgenommene Umformung die Lösungsmenge verändert wurde, die zweite Erklärung argumentiert mit der unzulässigen Division durch null (BIFIE 2016b: 3). 47 Prozent der Schülerinnen und Schüler haben einen dieser beiden Lösungswege gewählt und die Aufgabe im offenen Antwortformat richtig bearbeitet. Die restlichen Antworten teilen sich in sechs Kategorien auf.

3. Kategorien

Kategorie I beinhaltet all jene Antworten, die die Antwort mit der Unzulässigkeit der konkreten Rechnung $0:x$ begründen. Die Erklärung, die Gleichung müsse mittels Wurzelziehen gelöst werden, spiegelt sich in *Kategorie II* wieder. Alle Antworten, die *Kategorie III* zuzurechnen sind, stoßen sich daran, dass sich die rechte Seite der Gleichung nicht ändert. Die *vierte Kategorie (IV)* enthält schließlich jene Antworten, die zwar in die richtige Richtung gehen, per se also keine falsche Aussage enthalten, die Begründung aber nicht ausreicht um die Aufgabenstellung zu erfüllen. Dazu zählen insbesondere auch

Bearbeitungen, die sich nicht, wie in der Aufgabenstellung gefordert, konkret auf das Beispiel beziehen. Zu diesen vier Kategorien kommen noch die *Restkategorie (RK)* sowie die *Kategorie Nicht Gemacht (NG)*. Zusammen mit der *Kategorie der korrekt bearbeiteten Performanzen* ergibt sich Statistik 5.

4. Auswertung

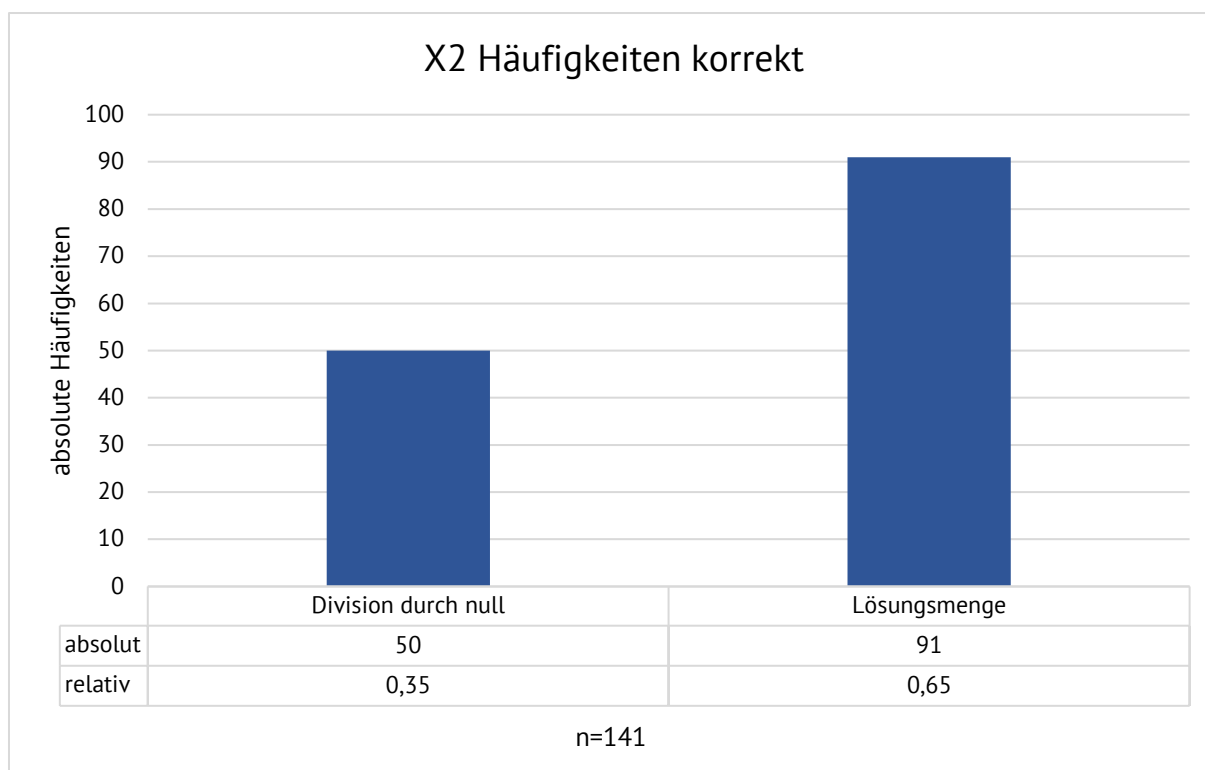


Statistik 5: X2 Grundgesamtheit: alle Performanzen

Laut Statistik 5 enthält Kategorie IV die meisten Entsprechungen unter den falsch gelösten Aufgaben. 20 Prozent aller Maturantinnen und Maturanten sind in der Beantwortung dieser Aufgabe zu wenig konkret gewesen und haben so die Aufgabenstellung nicht erfüllt. 44 Schülerinnen und Schüler haben sich für Antworten entschieden, die Kategorie III zuzurechnen sind, während Kategorie I und II mit jeweils 10 Entsprechungen am Ende des Spektrums zu finden sind. Gemeinsam mit der *Restkategorie* und der *Kategorie Nicht Gemacht* ergibt sich ein schöner Überblick über die verschiedenen Lösungsstrategien für Aufgabe 2.

5. Besonderheiten

Wie unter 2. *Beschreibung* bereits erwähnt, gibt die Lösungserwartung nur bei Aufgabe 2 zwei mögliche Lösungswege an. Wenn diese Unterscheidung in die Auswertung hineingenommen wird, lassen sich die korrekten Antworten noch einmal in zwei Kategorien einteilen. Dies wird in Statistik 6 dargestellt:



Statistik 6: X2 Grundgesamtheit: korrekte Bearbeitungen

Weit mehr Schülerinnen und Schüler (in absoluten Zahlen: 41) haben mit einer Veränderung der Lösungsmenge argumentiert. Statistik 6 zeigt, dass 35 Prozent der korrekten Antworten auf die Begründung der unzulässigen Division durch null entfielen.

4.3 Aufgabe 3

Treibstoffkosten

Der durchschnittliche Treibstoffverbrauch eines PKW beträgt y Liter pro 100 km Fahrtstrecke. Die Kosten für den Treibstoff betragen a Euro pro Liter.

Aufgabenstellung:

Geben Sie einen Term an, der die durchschnittlichen Treibstoffkosten K (in Euro) für eine Fahrtstrecke von x km beschreibt!

$K =$ _____

1. Grundkompetenz: AG 2.1

Einfache Terme und Formeln aufstellen, umformen und im Kontext deuten können

2. Beschreibung

Aufgabe 3 (X3) stellt sich als eine der Aufgaben dar, für die es am schwierigsten war, Kategorien zu finden. Bei 160 richtigen Antworten ($K = x \cdot \frac{y}{100} \cdot a$) und somit einer Lösungsquote von 53 Prozent bleiben noch 141 zu kategorisierende Performanzen. Dass dies nur mit Hilfe von 20 Kategorien möglich sein würde, entspricht nicht den Erwartungen an ein übersichtliches System. Allerdings wird bei der Kategorisierung der Lösungen zu dieser Aufgabe im halboffenen Antwortformat gerade Wert darauf gelegt, die große Vielfalt in den Antworten zu zeigen. Mit einer gröberen Einteilung wäre der Informationsverlust dahingehend bereits zu groß.

3. Kategorien

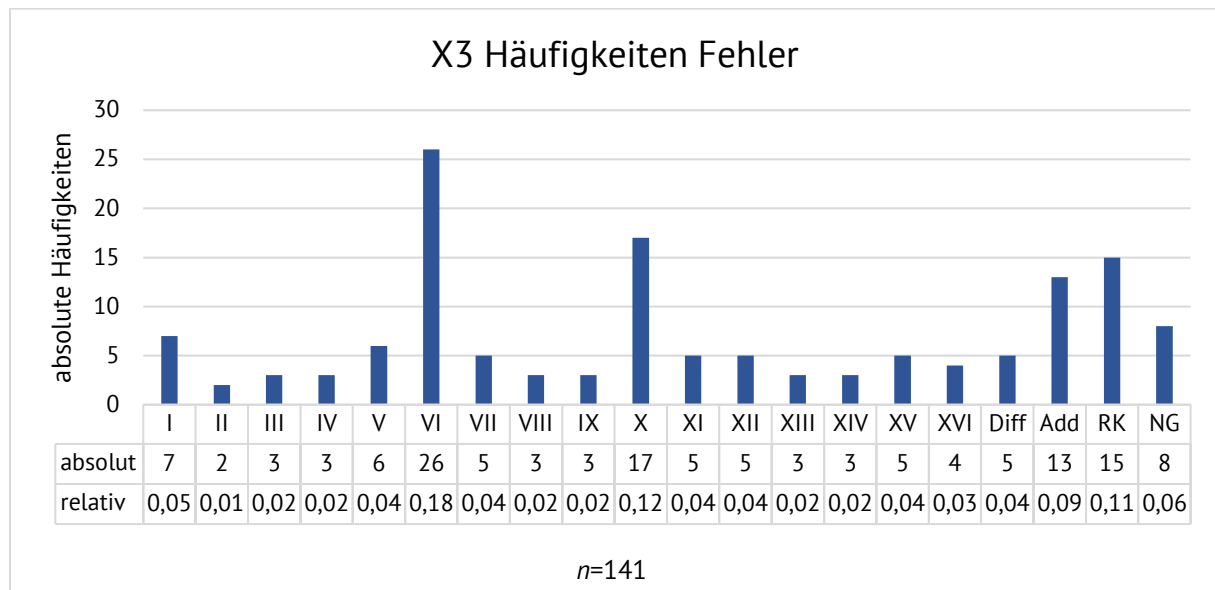
Nur drei Kategorien umfassen mehrere leicht unterschiedliche Antworten. Zum einen ist dies die *Kategorie Diff*, die alle Antworten zusammenfasst, die einen Differenzterm $(\frac{f(y)-f(a)}{y-a})$ beinhalten. Die *Kategorie Add* enthält all jene Performanzen, die einen additiven Term als Lösung machen. Die *Kategorie RK (Restkategorie)* umfasst alle falschen Antworten, die nur einmal aufgetreten sind. Die übrigen *Kategorien I-XVI* werden jeweils einem Term zugeteilt (siehe Tabelle 12). So entspricht beispielsweise *Kategorie I* dem Lösungsterm $\frac{a \cdot y}{100 \cdot x}$. Mit der *Kategorie NG, der nicht gemachten Aufgaben*, ergeben sich schlussendlich 20 Kategorien.

Kategorie	Beschreibung
I	$\frac{a \cdot y}{100 \cdot x}$
II	$\frac{100 \cdot y}{a \cdot x}$
III	$\frac{x \cdot y}{100 \cdot a}$
IV	$\frac{100 \cdot a \cdot x}{y}$
V	$\frac{100 \cdot a \cdot y}{x}$
VI	$\frac{a \cdot y}{x}$
VII	$\frac{a \cdot x}{y}$
VIII	$\frac{x \cdot y}{a}$
IX	$\frac{a}{x}$
X	$\frac{a \cdot y}{100}$
XI	$\frac{a \cdot x}{100}$
XII	$\frac{x \cdot y}{100}$
XIII	$100 \cdot a \cdot x \cdot y$
XIV	$a \cdot x \cdot y$
XV	$a \cdot x$
XVI	$a \cdot y$
Diff	Differenzterm
Add	additiver Term
RK	Restkategorie
NG	Nicht Gemacht

Tabelle 12: X3 Kategoriensystem

So zeigt Tabelle 12 eindrücklich, wie groß die Vielfalt der Fehlinterpretationen eines verbalen Textes ist beziehungsweise welche Fehler bei der Übersetzung in die Sprache der Algebra passieren können.

4. Auswertung



Statistik 7: X3 Grundgesamtheit: fehlerhafte Antworten

In Statistik 7 sind die Häufigkeiten der Entsprechungen zu den einzelnen Kategorien im Bezug auf die Grundgesamtheit der fehlerhaften Antworten zu sehen. Dabei hat jede Antwort mit mehr als einer Entsprechung eine eigene Kategorie erhalten. Antworten, die nur einmal aufgetreten sind, finden sich in der *Restkategorie (RK)* wieder. Die Kategorien *Diff* und *Add* wurden erstellt um eine übermäßig große *Restkategorie* zu vermeiden. Auch innerhalb dieser Kategorien gibt es beinahe ausschließlich Unikate, im Unterschied zur *Restkategorie*, immer noch mit einer, der jeweiligen Kategorie entsprechenden Gemeinsamkeit.

Durch das Festhalten an den einzelnen Kategorien zeigt sich die große Vielfalt in den Antworten (Tabelle 12). Auch wenn Kategorie VI mit 18 Prozent aller nicht korrekten Antworten überwiegt, kann doch nicht von einem vorwiegenden Fehler, wie es ihn bei Aufgabe 19 gibt, gesprochen werden. Mit Kategorie X, der zweithäufigsten Kategorie, gibt es eine Gemeinsamkeit: Bei beiden Kategorien steht der Ausdruck $a \cdot y$ im Zähler.

Der Ausdruck $\frac{a \cdot y}{x}$ aus Kategorie VI berechnet im Zähler die Kosten für 100 km Strecke und dividiert diese durch die tatsächlich gefahrene Strecke. Kategorie X enthält denselben Zähler dividiert aber durch 100. Dieser Ausdruck entspricht also den Kosten für einen Kilometer Fahrtstrecke – es fehlt die Berücksichtigung der gesamten Strecke (x). Ähnlich verhält es sich bei Kategorie XII: Hier wird durch den Term die benötigte Literanzahl für die gesamte Strecke ausgedrückt – dabei werden allerdings die Kosten pro Kilometer (a) vernachlässigt.

4.4 Aufgabe 4

Quadratische Gleichung

Gegeben ist die quadratische Gleichung $x^2 + p \cdot x - 12 = 0$.

Aufgabenstellung:

Bestimmen Sie denjenigen Wert für p , für den die Gleichung die Lösungsmenge $L = \{-2; 6\}$ hat!

1. Grundkompetenz: AG 2.3

Quadratische Gleichungen in einer Variablen umformen/lösen, über Lösungsfälle Bescheid wissen, Lösungen und Lösungsfälle (auch geometrisch) deuten können

2. Beschreibung

Mit einer durchschnittlichen Lösungsquote von 0,82 gehört die Aufgabe 4 (X4) des Haupttermins der ssRP Mathematik 2016 zu den Aufgaben mit den höheren Lösungsquoten. Innerhalb des Bereichs *Algebra und Geometrie* handelt es sich sogar um die Aufgabe mit der höchsten Lösungsquote. Mithilfe des Satzes von Vieta lässt sich die richtige Lösung $p = -4$ berechnen (offenes Aufgabenformat). Die hohe Anzahl an richtigen Antworten sowie auch eine hohe Anzahl an Performanzen, die die Aufgabe nicht bearbeitet haben, bedingen Kategorien mit einer geringen Zahl an Entsprechungen.

3. Kategorien

Insgesamt haben sich, abgesehen von der *Restkategorie (RK)*, den *nicht bearbeiteten Aufgaben (NG)* und den *richtigen Antworten*, drei Kategorien herausgebildet. Die *erste Kategorie (I)* beschreibt all jene Bearbeitungen, die versucht hatten, die Aufgabe mithilfe von Ausdrücken wie

$$x^2 + p \cdot x - 12 = x_1 \text{ bzw. } x_2$$

oder

$$x_1^2 + p \cdot x_1 - 12 = x_2$$

zu lösen, wobei $x_1 = -2$ und $x_2 = 6$. Ein Beispiel dazu ist in Abbildung 16 zu sehen.

Aufgabenstellung:

Bestimmen Sie denjenigen Wert für p , für den die Gleichung die Lösungsmenge $L = \{-2; 6\}$ hat!

$$\begin{aligned}
 (-2)^2 + p \cdot (-2) - 12 &= 6 \\
 4 - 2p - 12 &= 6 \\
 -2p - 8 &= 6 \\
 -2p &= \cancel{14} 14 \\
 p &= -7
 \end{aligned}$$

Abbildung 16: X4 Beispiel Kategorie I

Kategorie II beinhaltet all jene Performanzen, die versucht haben einen Lösungsweg mithilfe einer falschen Diskriminante zu finden. Ausgehend von

$$-\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - 12}$$

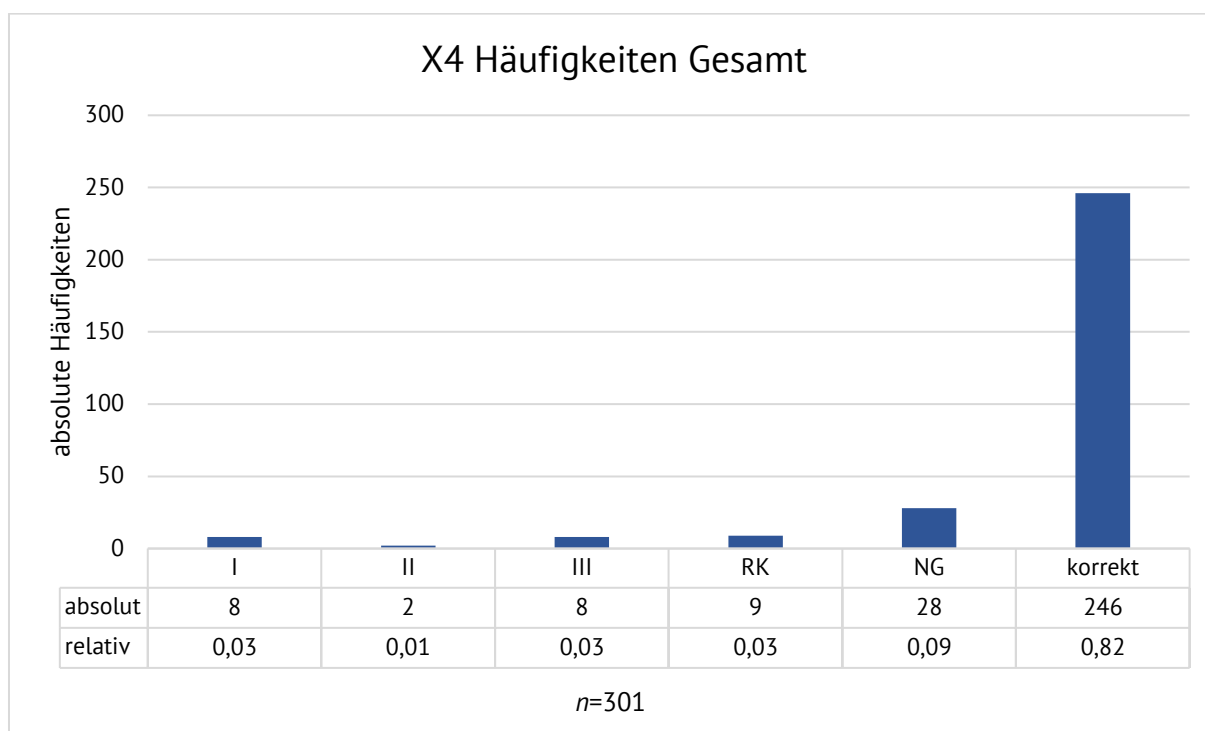
und der Forderung, dass

$$\frac{p^2}{4} - 12 > 0$$

führt dies in den vorliegenden Fällen zum Ergebnis $p > 4 \cdot \sqrt{3}$.

In der letzten verbleibenden *Kategorie III* werden die Ergebnisse mit Vorzeichenfehler zusammengefasst. Performanzen dieser Kategorie haben also das Ergebnis $p = 4$.

4. Auswertung



Statistik 8: X4 Grundgesamtheit: alle Performanzen

Statistik 8 zeigt: 28 Schülerinnen und Schüler (9 Prozent) haben die vierte Aufgabe nicht bearbeitet. Abzüglich der 246 korrekten Antworten ergeben sich so nur mehr 27 Performanzen, die kategorisiert werden können. Von diesen 27 Arbeiten entfällt wiederum ein Drittel auf die *Restkategorie*. Damit entfallen auf die einzelnen verbleibenden Kategorien weniger Entsprechungen als auf die *Restkategorie*. Innerhalb der *Restkategorie* erweist sich jede Antwort als Unikum, dies ist auch der Grund wieso für die Kategorie II, die nur zwei Entsprechungen beinhaltet, eine eigene Kategorie gebildet wurde. Bei dieser Aufgabe war die *Restkategorie* jenen Antworten vorbehalten, die nur einmal auftraten.

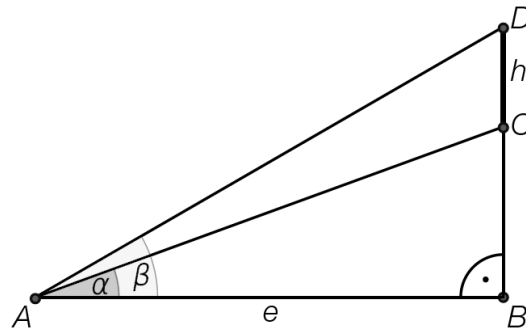
4.5 Aufgabe 5

Aufgabe 5 wurde unter 3.3 *Exemplarische Auswertung* detailliert beschrieben.

4.6 Aufgabe 6

Vermessung einer unzugänglichen Steilwand

Ein Steilwandstück CD mit der Höhe $h = \overline{CD}$ ist unzugänglich. Um h bestimmen zu können, werden die Entfernung $e = 6$ Meter und zwei Winkel $\alpha = 24^\circ$ und $\beta = 38^\circ$ gemessen. Der Sachverhalt wird durch die nachstehende (nicht maßstabgetreue) Abbildung veranschaulicht.



Aufgabenstellung:

Berechnen Sie die Höhe h des unzugänglichen Steilwandstücks in Metern!

1. Grundkompetenz: AG 4.1

Definitionen von *Sinus*, *Cosinus* und *Tangens* im rechtwinkligen Dreieck kennen und zur Auflösung rechtwinkliger Dreiecke einsetzen können

2. Beschreibung

Die letzte Aufgabe aus dem Bereich *Algebra und Geometrie* beschäftigt sich mit den trigonometrischen Funktionen. Mithilfe dieser soll die Länge eines Steilwandstückes berechnet werden. Die durchschnittliche Lösungsquote dabei lag bei 0,74. Damit handelt es sich bei Aufgabe 6 (X6) um die, nach der vierten Aufgabe, am zweithäufigsten korrekt gelöste Aufgabe aus dem ersten Teilbereich. Das Aufgabenformat ist offen.

Für die Berechnung der Länge h wird in der Lösungserwartung (BIFIE 2016b: 7) die Berechnung über die Tangensfunktion als eine mögliche Vorgehensweise vorgeschlagen.

$$\tan(\alpha) = \frac{\overline{BC}}{e} \Rightarrow \overline{BC} \approx 2,67 \text{ m}$$

$$\tan(\beta) = \frac{\overline{BD}}{e} \Rightarrow \overline{BD} \approx 4,69 \text{ m}$$

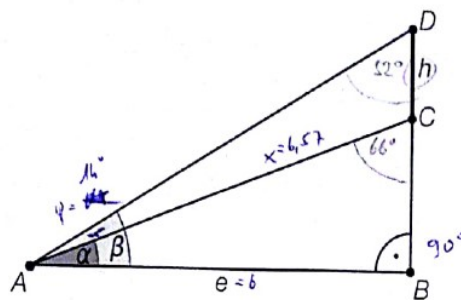
$$h = \overline{BD} - \overline{BC} \approx 2,02 \text{ m}$$

Der Lösungsschlüssel weiter:

Ein Punkt für die richtige Lösung, wobei die Einheit ,m' nicht angegeben sein muss. Die Aufgabe ist auch dann als richtig gelöst zu werten, wenn bei korrektem Ansatz das Ergebnis aufgrund eines Rechenfehlers nicht richtig ist. Toleranzintervall: $[2\text{ m}; 2,1\text{ m}]$

3. Kategorien

Fehlerkategorie I umfasst all jene Antworten, die Formeln für Winkelfunktionen in rechtwinkligen Dreiecken auf nichtrechtwinklige Dreiecke angewendet haben.



Aufgabenstellung:

Berechnen Sie die Höhe h des unzugänglichen Steilwandstücks in Metern!

$$\begin{aligned}\cos \alpha &= \frac{6}{x} \\ \cos \alpha \cdot x &= 6 \\ 0,913 \cdot x &= 6 \\ x &= 6,57\end{aligned}$$

$$\tan \varphi = \frac{h}{6,57}$$

$$0,249 = \frac{h}{6,57}$$

$$\underline{h = 1,638\text{ m}}$$

Abbildung 17: X6 Kategorie I

Wie in Abbildung 17 zu erkennen, wurde die erste Länge (hier x) korrekt berechnet. Für die Berechnung der zweiten Länge wurde die Formel $\tan(\varphi) = \frac{\text{Gegenkathete}}{\text{Ankathete}}$ auf das nichtrechtwinklige Dreieck $\triangle ACD$ angewendet.

Die zweite Kategorie (II) entsteht aus den Performanzen, die eine falsche Formel verwendet haben. Dazu gehören Ausdrücke der Form $\sin(\varphi) = \frac{\text{Ankathete}}{\text{Hypothenuse}}$ oder $\cos(\varphi) = \frac{\text{Ankathete}}{\text{Gegenkathete}}$. Auch Antworten wie in Abbildung 18, bei der trotz der ursprünglichen

korrekten Formel im Weiteren mit dem Arcustangens gerechnet wird, fallen unter diese Kategorie.

Aufgabenstellung:

Berechnen Sie die Höhe h des unzugänglichen Steilwandstücks in Metern!

$$\begin{aligned} \tan \alpha &= \frac{BC}{6} \\ 6 \cdot \tan \alpha &= BC = 525,684 \text{ m} \\ \tan \beta &= \frac{BC+h}{6} \\ (6 \tan \beta) - BC &= h \\ h &= 5,27104 \text{ m} \end{aligned}$$

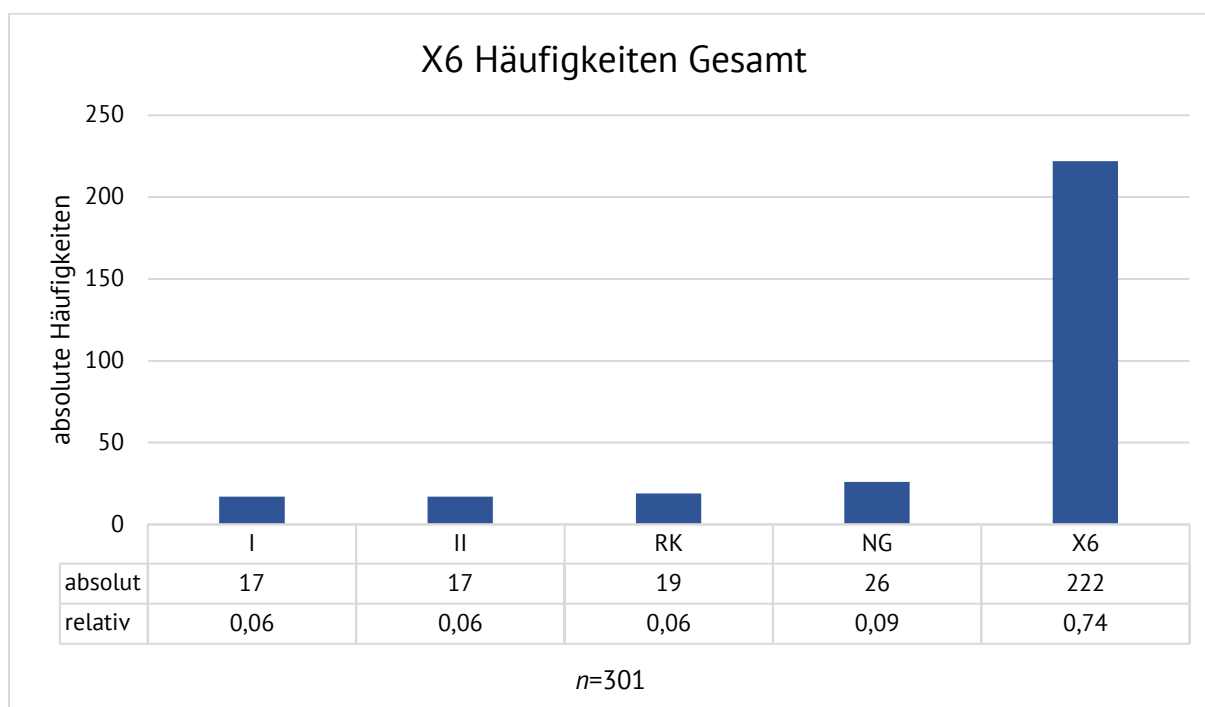
$$\begin{aligned} \tan \alpha &= \frac{BC}{6} \\ BC &= 2,671572712 \\ \tan \beta &= \frac{BC+h}{6} \end{aligned}$$

Abbildung 18: X6 Kategorie II

Antworten wie in Abbildung 18 traten innerhalb dieser Kategorie tendenziell etwas öfter auf, allerdings insgesamt zu wenig oft, um eine eigene Kategorie zu erstellen. So werden auch die Antworten mit falsch verwendetem Arcustangens in Kategorie II subsumiert. Ganz nach dem Wesen der Kategorienbildung geht so zwar eine Detailinformation verloren, der Überblick über die Aufgabe wird allerdings einfacher.

Zusätzlich zu diesen beiden Kategorien werden in Statistik 9 noch die *Kategorie der nicht gemachten Performanzen (NG)*, die *Restkategorie (RK)* sowie die *korrekt gelösten Aufgaben* dargestellt.

4. Auswertung



Statistik 9: X6 Grundgesamtheit: alle Performanzen

Abgesehen von den korrekten Antworten bildet die *Kategorie der nicht gemachten Aufgaben* mit 26 Entsprechungen die größte Kategorie. Statistik 9 zeigt auch, dass die *Restkategorie* größer als die beiden konkreten Kategorien I und II ist. 19 Performanzen lassen sich weder Kategorie I noch Kategorie II zuordnen, aber auch unter ihnen gibt es keine Häufung gemeinsamer Merkmale, die eine eigene Kategorie rechtfertigen würden.

5. Besonderheiten

Wie bereits bei Kategorie II erwähnt, kommen auch Ausdrücke die den Arcustangens enthalten, dort immer wieder einmal vor. Aber auch der Sinus- und Cosinussatz tauchen auf und werden falsch eingesetzt, allerdings in einer zahlenmäßig zu geringen Ausprägung, um bei Kategorien und Auswertung eine Rolle zu spielen. Diese drei verschiedenen Lösungsstrategien haben jedoch eines gemeinsam: Wird der Grundkompetenzenkatalog zur standardisierten schriftlichen Reifeprüfung Mathematik (BIFIE 2013a) etwas genauer betrachtet, wird schnell klar, dass gerade die oben erwähnten Werkzeuge gar nicht Stoff des grundkompetenzenabfragenden Teil 1 sind. Die bereits erläuterte Grundkompetenz AG 4.1 verlangt explizit nur „Sinus, Cosinus und Tangens im rechtwinkligen Dreieck“ und eine Anmerkung macht klar: „Sinus- und Cosinussatz werden dabei [Auflösung rechtwinkliger Dreiecke, Anm.] nicht benötigt“ (BIFIE 2013a: 8).

4.7 Aufgabe 19

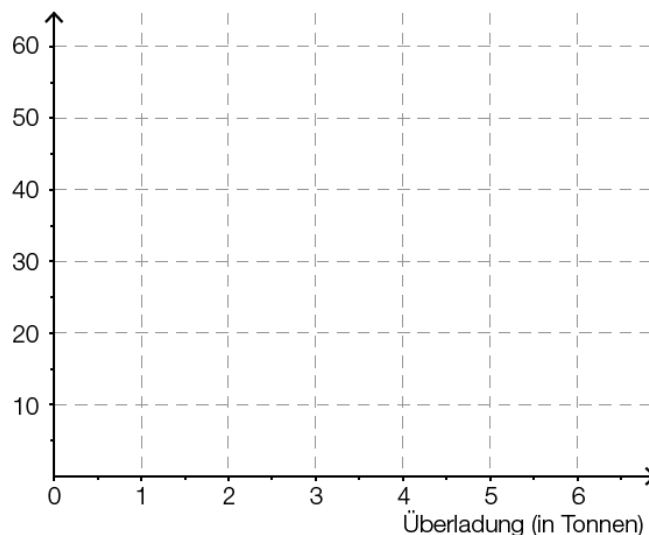
Beladung von LKW

Bei einer Verkehrskontrolle wurde die Beladung von LKW überprüft. 140 der überprüften LKW waren überladen. Details der Kontrolle sind in der nachstehenden Tabelle zusammengefasst.

Überladung $\ddot{U}$ in Tonnen	$\ddot{U} < 1\text{t}$	$1\text{t} \leq \ddot{U} < 3\text{t}$	$3\text{t} \leq \ddot{U} < 6\text{t}$
Anzahl der LKW	30	50	60

Aufgabenstellung:

Stellen Sie die Daten der obigen Tabelle durch ein Histogramm dar! Dabei sollen die absoluten Häufigkeiten als Flächeninhalte von Rechtecken abgebildet werden.



1. Grundkompetenz: WS 1.2

Tabellen und einfache statistische Grafiken erstellen, zwischen Darstellungsformen wechseln können

2. Beschreibung

Die erste Aufgabe aus dem Gebiet *Wahrscheinlichkeit und Statistik* des Haupttermins 2016 verlangt die Erstellung eines Histogramms. Mit einer Lösungsquote von 0,27 handelt es sich um die Aufgabe mit der geringsten durchschnittlichen Lösungsquote aller untersuchten Aufgaben. Anders als bei Säulendiagrammen, bei denen die Höhe der Säulen das Datum anzeigt, ist bei Histogrammen der Flächeninhalt von Rechtecken das ausschlaggebende Kriterium. Diese Rechtecke sollen in das vorgegebene Koordinatensystem eingezeichnet werden, es handelt sich also um eine Konstruktionsaufgabe, mit der in Abbildung 19 (X19) abgebildeten Lösungserwartung (BIFIE 2016b: 20).

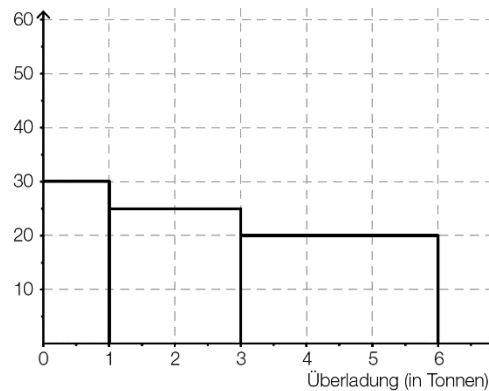


Abbildung 19: X19 Lösungserwartung

Im Zuge der Kategorienfindung stellte sich rasch heraus, dass es eine sehr große Kategorie geben würde. Nicht abzusehen war jedoch, dass keine weiteren Kategorien gebildet werden können. Die *Restkategorie (RK)* ist zwar verhältnismäßig groß, allerdings lässt sich innerhalb dieser anzahlmäßig keine Unterkategorie finden, die groß genug wäre um eine eigene Kategorie rechtfertigen zu können.

3. Kategorien

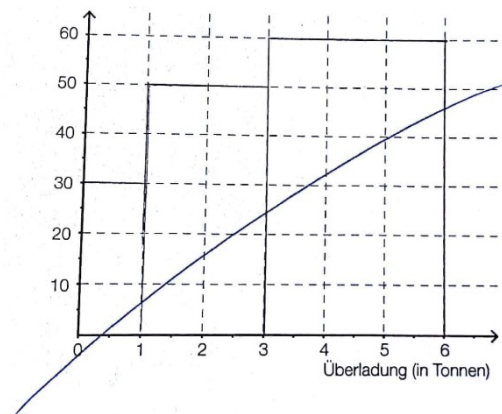
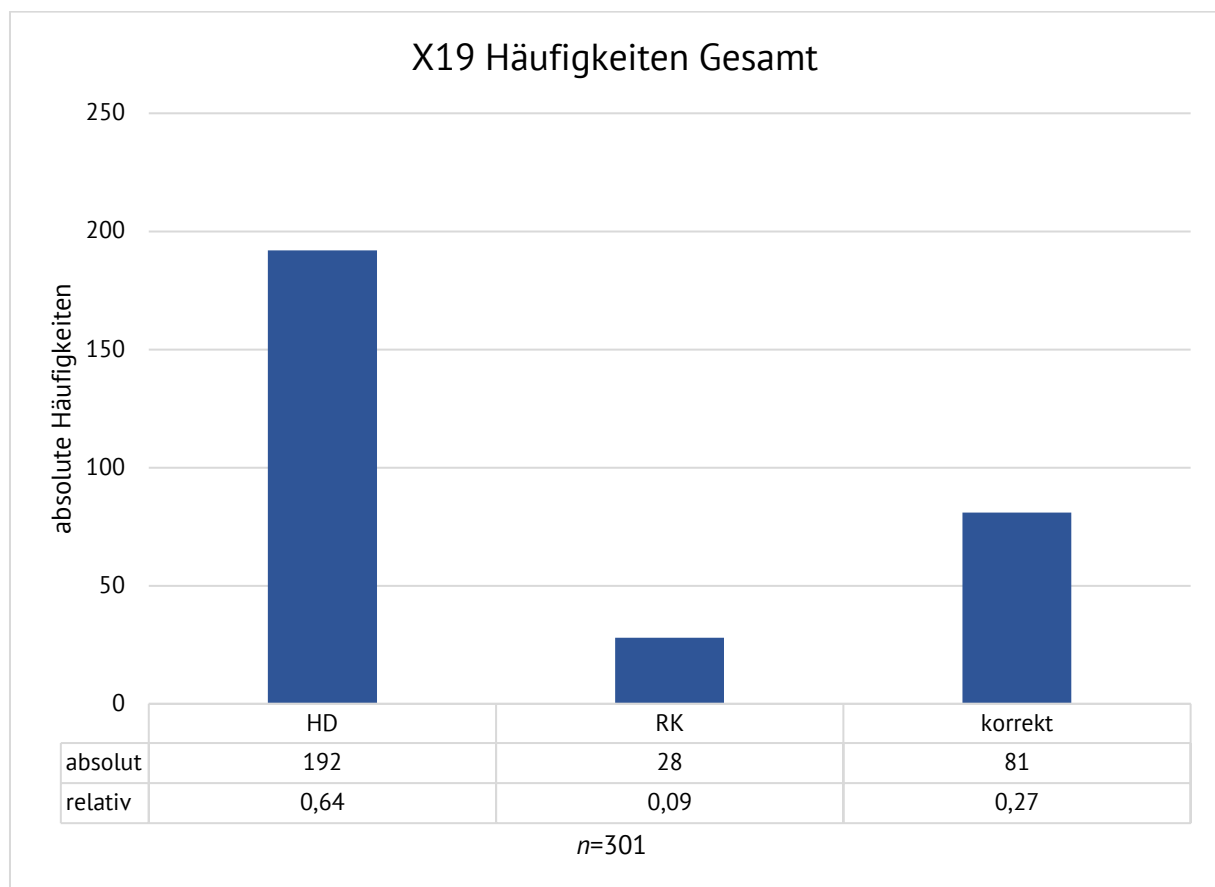


Abbildung 20: X19 Kategorie HD

Wie in Abbildung 20 zu sehen, wurde in der *Kategorie HD* die Häufigkeitsdichte nicht beachtet. Das heißt, dass die absoluten Werte für die jeweiligen Bereiche entlang der senkrechten Achse abgetragen wurden, ohne auf die dadurch entstehende Fläche zu achten. Des Weiteren werden in der folgenden Grafik zur Auswertung (Statistik 10) eine *Restkategorie RK* sowie die *Kategorie der korrekten Antworten* geführt.

4. Auswertung



Statistik 10: X19 Grundgesamtheit: alle Performanzen

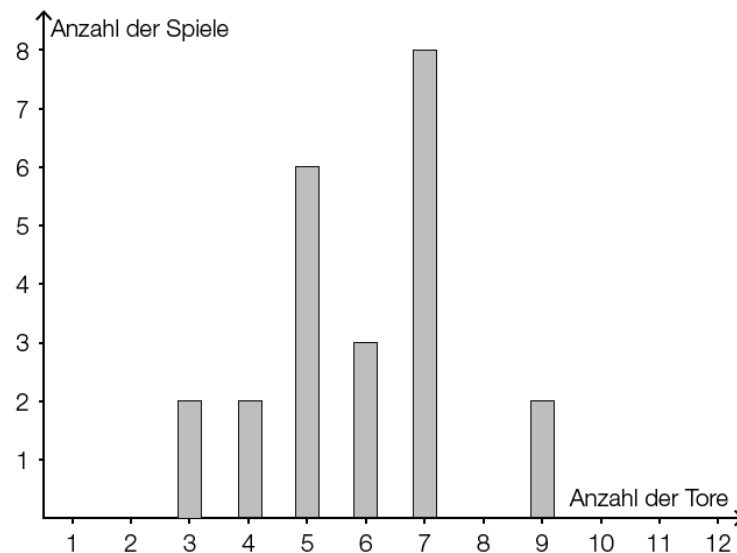
Sowohl aus den Häufigkeitstabellen und dem Säulendiagramm aus Statistik 10 wird klar welche Bedeutung diese eine Kategorie *HD* besitzt. 64 Prozent der Maturantinnen und Maturanten und damit fast zwei Drittel haben denselben Fehler gemacht. Demgegenüber stehen 81 Schülerinnen und Schüler (27 Prozent), die die Aufgabe korrekt lösen konnten und weitere 28 Performanzen, die in keine Kategorie eingeordnet werden konnten.

Damit stellt die Kategorie *HD* auch die Fehlerkategorie mit den meisten Entsprechungen unter allen Aufgaben dar. Zusammen mit dem Umstand, dass keine weiteren Kategorien beschrieben werden konnten, kann hier bei Aufgabe 19 vom deutlichsten aller Fehlerbilder gesprochen werden. Auch die vollständige Bearbeitung – alle 301 Schülerinnen und Schüler haben diese Aufgabe beantwortet – unterstreicht diese Tatsache weiter.

4.8 Aufgabe 20

Eishockeytore

In der österreichischen Eishockeyliga werden die Ergebnisse aller Spiele statistisch ausgewertet. In der Saison 2012/13 wurde über einen bestimmten Zeitraum erfasst, in wie vielen Spielen jeweils eine bestimmte Anzahl an Toren erzielt wurde. Das nachstehende Säulendiagramm stellt das Ergebnis dieser Auswertung dar.



Aufgabenstellung:

Bestimmen Sie den Median der Datenliste, die dem Säulendiagramm zugrunde liegt!

1. Grundkompetenz: WS 1.3

Statistische Kennzahlen (absolute und relative Häufigkeiten; arithmetisches Mittel, Median, Modus, Quartile, Spannweite, empirische Varianz/Standardabweichung) im jeweiligen Kontext interpretieren können; die angeführten Kennzahlen für einfache Datensätze ermitteln können

2. Beschreibung

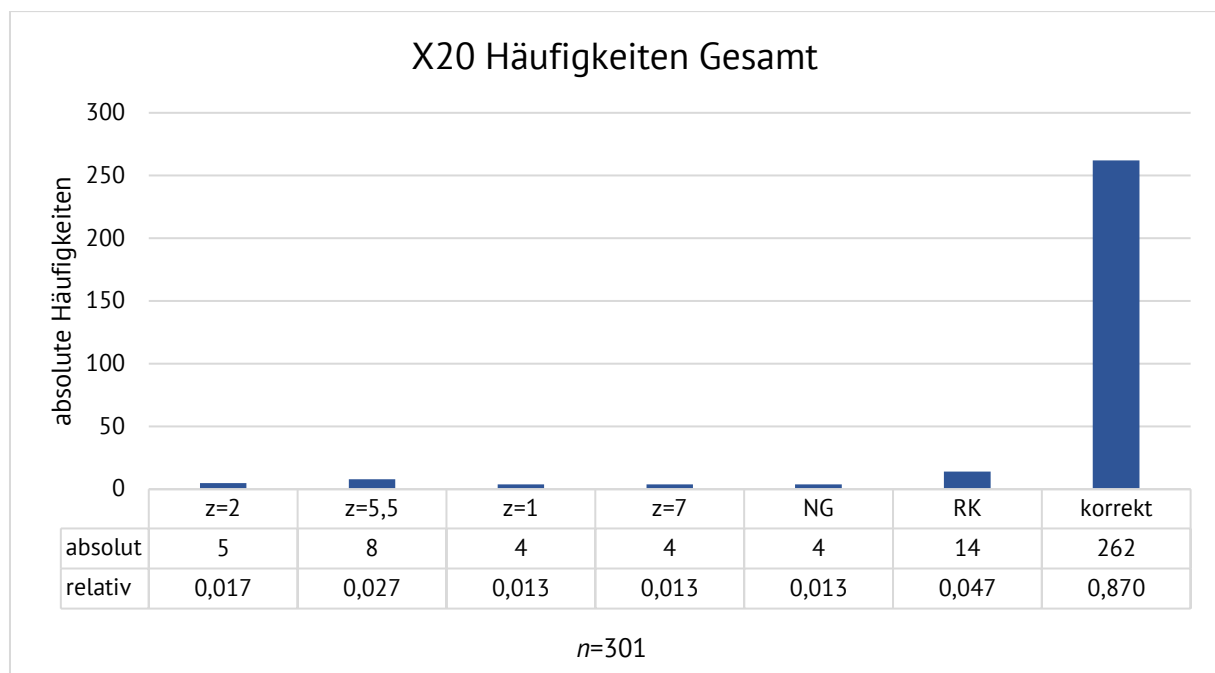
Um die Grundkompetenz *WS 1.3* abzu prüfen, wurde eine Aufgabe im offenen Format gewählt. Mit einer Lösungsquote von 0,87 stellt die Aufgabe 20 (X20) jene mit der vierthöchsten Lösungsquote der Teil-1-Aufgaben und jene mit der höchsten Lösungsquote der in der vorliegenden Arbeit behandelten Aufgaben dar. Von den in der Grundkompetenz genannten statistischen Kennzahlen soll hier der Median aus dem Säulendiagramm herausgelesen werden. Bei der gegebenen Datenliste lautet der Median $z = 6$.

Aufgrund der hohen Lösungsquote von 87 Prozent (262 Schülerinnen und Schüler), bleiben dementsprechend wenige falsche Antworten. Je nach Anzahl der Kategorien ist also eine sehr geringe Prozentzahl pro Kategorie zu erwarten. Bei anderen Aufgaben würde das unter Umständen zurecht zu einer Zusammenfassung zu einer *Restkategorie* führen. Hier muss die Grenze für die *Restkategorie* deutlich niedriger angesetzt werden, um die Unterschiede in den falschen Antworten überhaupt sichtbar machen zu können.

3. Kategorien

Die Kategorien werden so nah wie möglich an den Performanzen selbst gehalten. Folgerichtig werden die verschiedenen Antworten für den Median als Kategorien herangezogen und in weiterer Folge gezählt, wie oft ein bestimmter Wert als Median vorkommt. Aufgrund von fehlenden Aufzeichnungen ist es auch nicht möglich andere Aussagen zu treffen. Zu den *Kategorien*, denen jeweils ein falscher Medianwert zugeordnet wird, kommen noch die *Kategorie Nicht gemacht (NG)*, die *Restkategorie (RK)* sowie die *korrekten Antworten*.

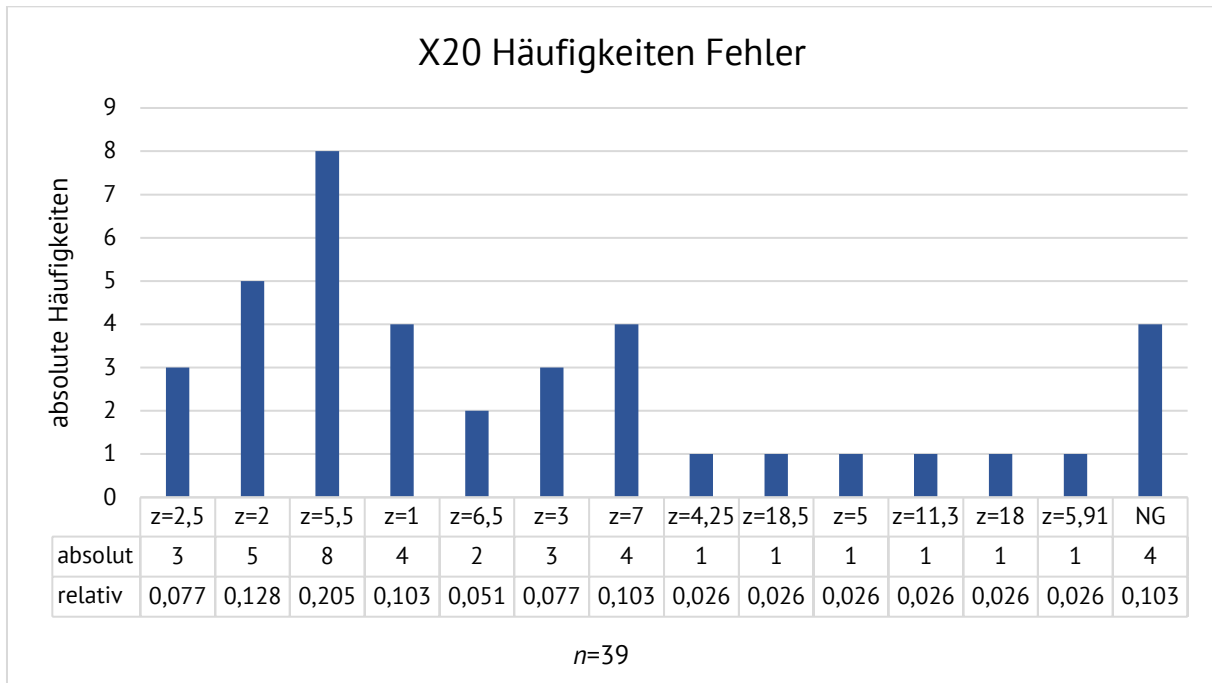
4. Auswertung



Statistik 11: X20 Grundgesamtheit: alle Performanzen

Bei einer derart großen Anzahl an korrekt bearbeiteten Performanzen kann mit einem Säulendiagramm, wie in Statistik 11, vor allem der enorme Unterschied zwischen der Kategorie der korrekten Antworten und den übrigen Kategorien sowie die minimalen Unterschiede der übrigen Kategorien unter sich gezeigt werden. Die Verteilung der Fehler

wird etwas klarer, wenn die Grundgesamtheit der fehlerhaften Arbeiten herangezogen wird (hier: $n=39$). In Statistik 12 wird keine *Restkategorie* geführt, so kann auch einmal die große Bandbreite an unterschiedlichen Antworten gezeigt werden.

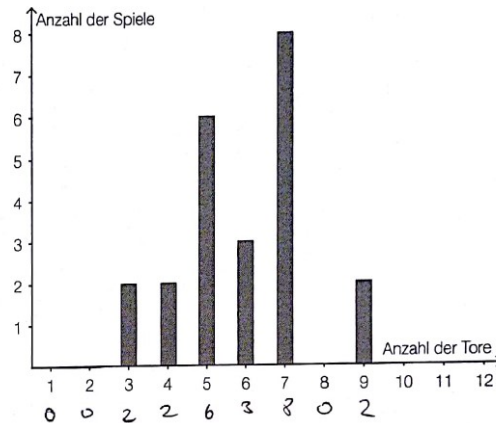


Statistik 12: X20 Grundgesamtheit: fehlerhafte Arbeiten

5. Besonderheiten

Ohne Aufzeichnungen zu den Aufgaben lassen sich keine anderen Kategorien bilden, wie die hier vorgestellten. Dies stellt jedoch in keiner Weise ein Hindernis dar, die so gefundenen Kategorien weiter zu untersuchen. Naheliegend ist zunächst den Modus zu untersuchen: Der am häufigsten falsch gewählte Median ist $z = 5,5$. Nach einer weiteren Untersuchung der Daten wird klar, dass innerhalb dieser acht Schülerinnen und Schüler fünf dieselbe Strategie verwendet haben. Diese fünf Maturantinnen und Maturanten haben die Datenliste 3,4,5,6,7,9 gewählt und aus dieser den Median gebildet. Diese Datenliste entspricht genau denjenigen Anzahlen von Toren, die in Spielen tatsächlich erzielt worden sind.

Bei erneuter Durchsicht der Daten fällt etwas anderes auf: Elf Schülerinnen und Schüler (und somit 28 Prozent der fehlerhaften Antworten) haben den Median der Anzahl der Spiele berechnet. Dabei haben sie eine verschiedene Anzahl von Nullen in die Datenliste zur Ermittlung des Medians aufgenommen und den Median daraufhin aus diesen Anzahlen ermittelt.



Aufgabenstellung:

Bestimmen Sie den Median der Datenliste, die dem Säulendiagramm zugrunde liegt!

~~0, 0, 2, 2, 6, 3, 8, 0, 2~~
 0, 0, 0, 2, 2, 2, 3, 6, 8;
 ↓
 2

~~M = 2~~

A = Der Median der Datenliste ist 2.

Abbildung 21: X20 Medianberechnung Spielanzahl

In Abbildung 21 wurden nur die Nullen vor der ersten Säule sowie die Null für die Anzahl der Spiele in denen acht Tore geschossen wurden ergänzt, dies führt zu folgenden Daten:

0, 0, 2, 2, 6, 3, 8, 0, 2

beziehungsweise geordnet

0, 0, 0, 2, 2, 2, 3, 6, 8.

Aus dieser Datenliste ergibt sich der Median $z = 2$.

Bei anderen Arbeiten wiederum wird nur die Null inmitten der Säulen, also für den Fall, dass acht Tore geschossen wurden berücksichtigt. Daraus resultiert ebenfalls ein Median von $z = 2$. Wurden auch noch die Nullen für den Fall, dass die Anzahl der Tore zehn, elf oder zwölf wäre, berücksichtigt, ergibt sich mit der Datenliste

0, 0, 2, 2, 6, 3, 8, 0, 2, 0, 0, 0

beziehungsweise geordnet

0, 0, 0, 0, 0, 0, 2, 2, 2, 3, 6, 8

der Median $z = 1$.

Wurden keine Nullen berücksichtigt, folgt bei dieser Berechnungsweise der Median $z = 2,5$. Anhand dieser Interpretationen ist schön zu sehen, dass scheinbar nicht zusammenhängende Ergebnisse doch auf einer gemeinsamen Grundidee fußen können.

4.9 Aufgabe 21

Zollkontrolle

Eine Gruppe von zehn Personen überquert eine Grenze zwischen zwei Staaten. Zwei Personen führen Schmuggelware mit sich. Beim Grenzübertritt werden drei Personen vom Zoll zufällig ausgewählt und kontrolliert.

Aufgabenstellung:

Berechnen Sie die Wahrscheinlichkeit, dass unter den drei kontrollierten Personen die beiden Schmuggler der Gruppe sind!

1. Grundkompetenz: WS 2.3

Wahrscheinlichkeit unter der Verwendung der Laplace-Annahme (Laplace-Wahrscheinlichkeit) berechnen und interpretieren können, Additionsregel und Multiplikationsregel anwenden und interpretieren können

Anmerkung: Die Multiplikationsregel kann unter Verwendung der kombinatorischen Grundlagen und der Anwendung der Laplace-Regel (auch) umgangen werden.

2. Beschreibung

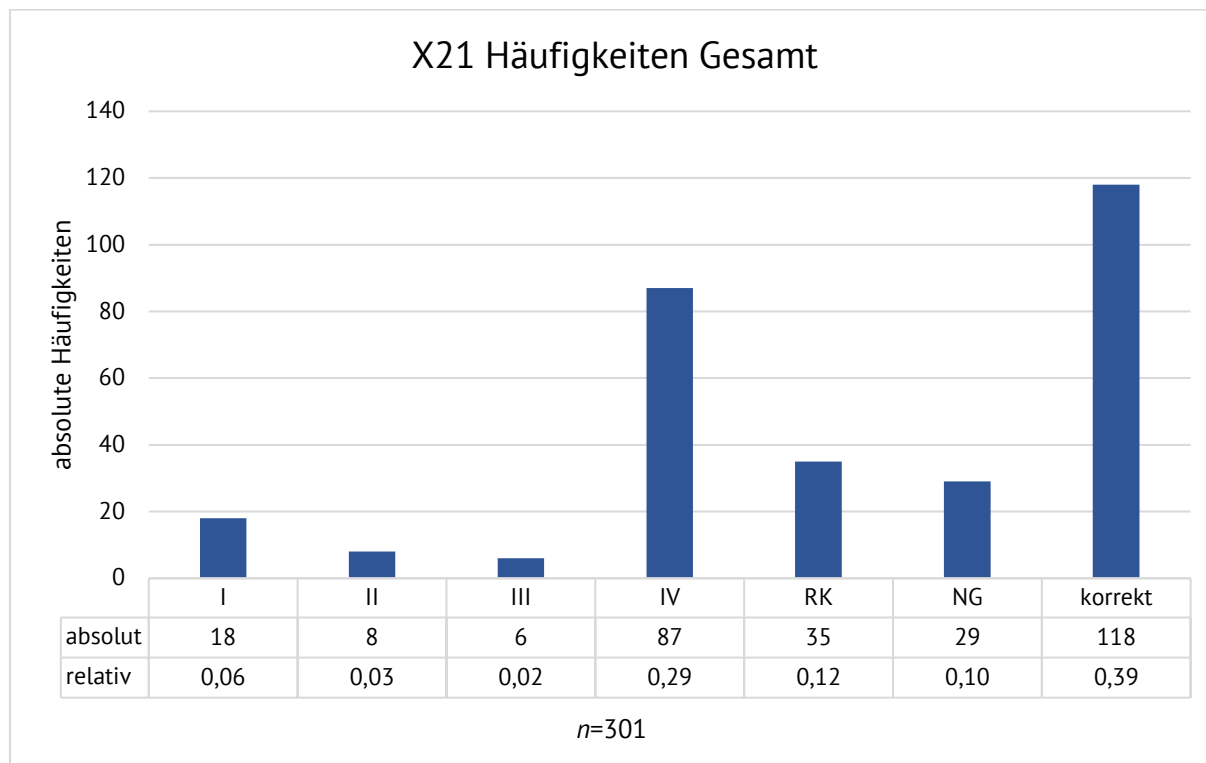
Für Aufgabe 21 (X21) zur Laplace-Wahrscheinlichkeit wird die Lösungserwartung $\frac{2}{10} \cdot \frac{1}{9} \cdot 3 = \frac{1}{15}$ angegeben (BIFIE 2016b: 22). Als Toleranzintervall wird [0,066; 0,07] beziehungsweise [6,6 %; 7 %] genannt. Diese Lösung haben 39 Prozent der Schülerinnen und Schüler angegeben. Damit weist die Aufgabe im offenen Antwortformat die zweitniedrigste durchschnittliche Lösungsquote aller untersuchten Aufgaben auf. Die 183 Performanzen, die keinen Punkt erlangen konnten, können in die folgenden Kategorien unterteilt werden.

3. Kategorien

18 Schülerinnen und Schüler haben bei ihrem Lösungsweg einen Weg zu wenig erkannt beziehungsweise die Reihenfolge nicht berücksichtigt. Dies bildet die *Kategorie I*. *Kategorie II* umfasst all jene Performanzen, die nicht mit den korrekten Anzahlen gerechnet haben. Innerhalb dieser Kategorie finden sich alle Arbeiten, die entweder einen Schmuggler zu viel oder zu wenig berücksichtigen, oder auch mit einer falschen Gesamtzahl an Personen rechnen. Schülerinnen und Schüler, die die Aufgabe *mit Zurücklegen* gelöst haben, werden in *Kategorie III* zusammengefasst. *Kategorie IV* enthält Performanzen, welche versucht haben, die Aufgabe über eine Binomialverteilung zu lösen. Zusammen mit der *Restkategorie (RK)*,

der *Kategorie der nicht gemachten Aufgaben (NG)* sowie der *korrekten Antworten* ergibt sich Statistik 13.

4. Auswertung



Statistik 13: X21 Grundgesamtheit: alle Performanzen

Bei einer derart großen *Restkategorie* wie bei der vorliegenden Aufgabenstellung stellt sich zurecht die Frage, ob die verhältnismäßig doch recht kleinen Kategorien II und III nicht auch der *Restkategorie* zugeordnet werden hätten können. Um nicht zu viel Information zu verlieren, wurde die Grenze für die *Restkategorie* aber bewusst niedrig gehalten. Schon ab vier Entsprechungen wurde eine neue Kategorie gebildet.

Die mit Abstand größte Fehlerkategorie stellt die Kategorie IV dar. 87 Schülerinnen und Schüler beziehungsweise 29 Prozent aller Angetretenen haben versucht die Aufgabe mithilfe einer Binomialverteilung zu lösen. An zweiter Stelle, abgesehen von *RK* und *NG*, folgt dann schon Kategorie I: 18 Maturantinnen und Maturanten haben die Reihenfolge nicht berücksichtigt.

5. Besonderheiten

Wie unter 1. *Grundkompetenz* erwähnt, kann die Multiplikationsregel auch umgangen werden. Dazu werden die kombinatorischen Grundlagen benötigt um die Wahrscheinlichkeit

nach Laplace als $\frac{\text{Günstige}}{\text{Mögliche}}$ darstellen zu können. Einige Schülerinnen und Schüler sind diesen Weg gegangen. Dabei konnten drei von sechs die richtige Lösung

$$\frac{\binom{2}{2} \cdot \binom{8}{1}}{\binom{10}{3}}$$

finden. Die korrekt gelösten Performanzen fallen natürlicherweise in die *Kategorie der korrekten Antworten*, die falsch bearbeiteten werden in der *Restkategorie* subsumiert.

4.10 Aufgabe 22

Wahrscheinlichkeitsverteilung

Der Wertebereich einer Zufallsvariablen X besteht aus den Werten x_1, x_2, x_3 .

Man kennt die Wahrscheinlichkeit $P(X = x_1) = 0,4$. Außerdem weiß man, dass x_3 doppelt so wahrscheinlich wie x_2 ist.

Aufgabenstellung:

Berechnen Sie $P(X = x_2)$ und $P(X = x_3)$!

$P(X = x_2) =$ _____

$P(X = x_3) =$ _____

1. Grundkompetenz: WS 3.1

Die Begriffe *Zufallsvariable*, (*Wahrscheinlichkeits*-)*Verteilung*, *Erwartungswert* und *Standardabweichung* verständig deuten und einsetzen können

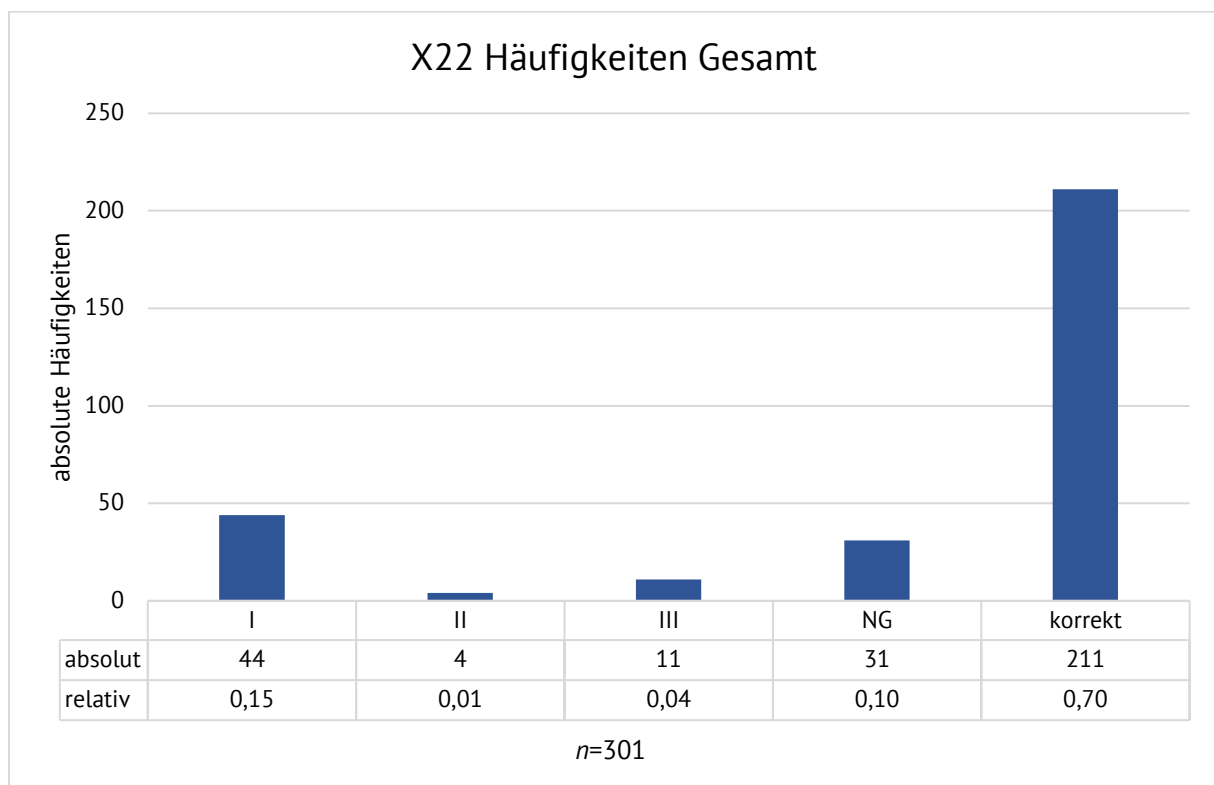
2. Beschreibung

211 Schülerinnen und Schülern und somit 70 Prozent der Angetretenen konnten die halboffene Aufgabe 22 (X22) korrekt lösen. Naturgemäß erfüllt die richtige Lösung $P(X = x_2) = 0,2$ und $P(X = x_3) = 0,4$ die Bedingung, $P(\Omega) = 1$, wobei $\Omega = \{x_1, x_2, x_3\}$. Nach diesem Aspekt wurde auch die Kategorisierung in Angriff genommen und grundlegend zwischen Antworten, die diese Eigenschaft beachtet haben und jenen die dies nicht getan haben, unterschieden.

3. Kategorien

Kategorie I enthält demnach all jene Performanzen, bei denen gilt $P(X = x_1) + P(X = x_2) + P(X = x_3) \neq 1$. Demnach gibt es auch eine *Kategorie II*, bei der zwar für jede Entsprechung $P(X = x_1) + P(X = x_2) + P(X = x_3) = 1$ gilt, die Aufgabe aber dennoch nicht korrekt gelöst wurde. Mithilfe dieser beiden Kategorien sind alle Fälle abgedeckt – könnte die Vermutung sein. Die Daten bedingen jedoch eine weitere *Kategorie III*. Diese enthält all jene Antworten, bei der die Summe der Wahrscheinlichkeiten über die Ergebnismenge (Ω) keinen numerischen Wert ergibt. Dabei wurde die Aufgabe in einem gegenseitigen Abhängigkeitsverhältnis gelöst. Als Antwort ergeben sich dann Ausdrücke wie $P(X = x_2) = \frac{x_3}{2}$ und $P(X = x_3) = 2 \cdot x_2$. Bleiben schlussendlich noch die üblichen Kategorien *Nicht Gemacht (NG)* und die *Kategorie der korrekten Antworten*.

4. Auswertung



Statistik 14: X22 Grundgesamtheit: alle Performanzen

Unter allen Performanzen haben 215 Schülerinnen und Schüler den korrekten Zusammenhang $P(\Omega) = 1$ erkannt, der für 211 Maturantinnen und Maturanten auch in der richtigen Antwort resultierte. 15 Prozent der Angetretenen haben die Problemstellung zwar auch numerisch gelöst, allerdings wurde der zuvor erwähnte Zusammenhang nicht berücksichtigt. Wie in Statistik 14 zu erkennen handelt es sich bei dieser Kategorie I, abgesehen von der *Kategorie Nicht Gemacht*, um die mit Abstand größte Fehlerkategorie. Viermal so viele Entsprechungen wie in Kategorie III und um den Faktor 11 mehr Entsprechungen als in Kategorie II sprechen eine klare Sprache.

4.11 Aufgabe 23

Verschiedenfarbige Kugeln

Auf einem Tisch steht eine Schachtel mit drei roten und zwölf schwarzen Kugeln. Nach dem Zufallsprinzip werden nacheinander drei Kugeln aus der Schachtel gezogen, wobei die gezogene Kugel jeweils wieder zurückgelegt wird.

Aufgabenstellung:

Gegeben ist der folgende Ausdruck:

$$3 \cdot 0,8^2 \cdot 0,2$$

Kreuzen Sie dasjenige Ereignis an, dessen Wahrscheinlichkeit durch diesen Ausdruck berechnet wird!

Es wird höchstens eine schwarze Kugel gezogen.	☐
Es werden genau zwei schwarze Kugeln gezogen.	☐
Es werden zwei rote Kugeln und eine schwarze Kugel gezogen.	☐
Es werden nur rote Kugeln gezogen.	☐
Es wird mindestens eine rote Kugel gezogen.	☐
Es wird keine rote Kugel gezogen.	☐

1. Grundkompetenz: WS 3.2

Binomialverteilung als Modell einer diskreten Verteilung kennen – Erwartungswert sowie Varianz/Standardabweichung binomialverteilter Zufallsgrößen ermitteln können, Wahrscheinlichkeitsverteilung binomialverteilter Zufallsgrößen angeben können, Arbeiten mit der Binomialverteilung in anwendungsorientierten Bereichen

2. Beschreibung

Bei Aufgabe 23 (X23) handelt es sich um die einzige Aufgabe des Formats *Multiple-Choice 1 aus 6* des Haupttermins 2016. 84,7 Prozent der Schülerinnen und Schüler setzten das Kreuz dabei an der richtigen zweiten Stelle. Mit dieser Lösungsquote weist die Aufgabe zur Binomialverteilung die zweithöchste Lösungsquote aus dem Bereich *Wahrscheinlichkeit und Statistik* auf.

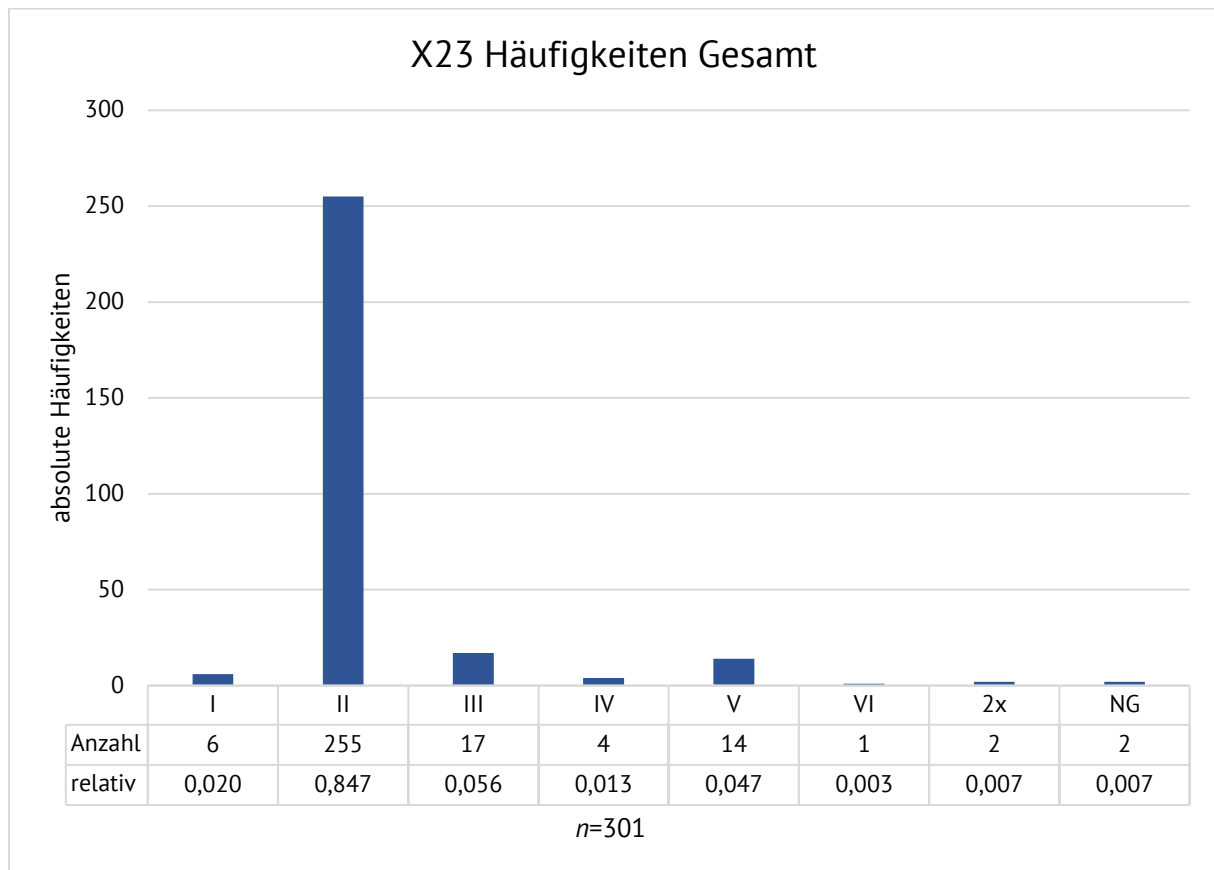
3. Kategorien

Die verschiedenen Antworten werden von oben nach unten mit den verschiedenen Kategorien identifiziert. Die erste Antwort entspricht folglich der Kategorie I, die letzte Antwort der Kategorie VI (Tabelle 13). Außerdem gibt es noch zwei weitere Kategorien: Zum einen wurden bei zwei Performanzen zwei Kreuze gemacht, diese beiden Arbeiten fallen in eine eigene Kategorie (2x). Zum anderen die *Kategorie Nicht gemacht (NG)*. Anders als bei Aufgabe 1 oder Aufgabe 24 können sowohl 2x als auch NG als eigene Kategorie geführt werden. Nachdem für die korrekte Bearbeitung der Aufgabe keine Antwortkombination notwendig ist, werden durch die Führung einer eigenen Kategorie für zwei Kreuze Mehrfachnennungen ausgeschlossen. Die Berücksichtigung der *Kategorie NG*, anders als bei Aufgabe 24 (bei dieser Aufgabe sind ebenfalls Kombinationen notwendig), führt dazu, dass mit allen Performanzen gearbeitet werden kann, die Grundgesamtheit also $n=301$ bleibt.

Kategorie	Beschreibung
I	Es wird höchstens eine schwarze Kugel gezogen.
II	Es werden genau zwei schwarze Kugeln gezogen.
III	Es werden zwei rote Kugeln und eine schwarze Kugel gezogen.
IV	Es werden nur rote Kugeln gezogen.
V	Es wird mindestens eine rote Kugel gezogen.
VI	Es wird keine rote Kugel gezogen.
2x	Es wurden zwei Kreuze gesetzt.
NG	Nicht gemacht

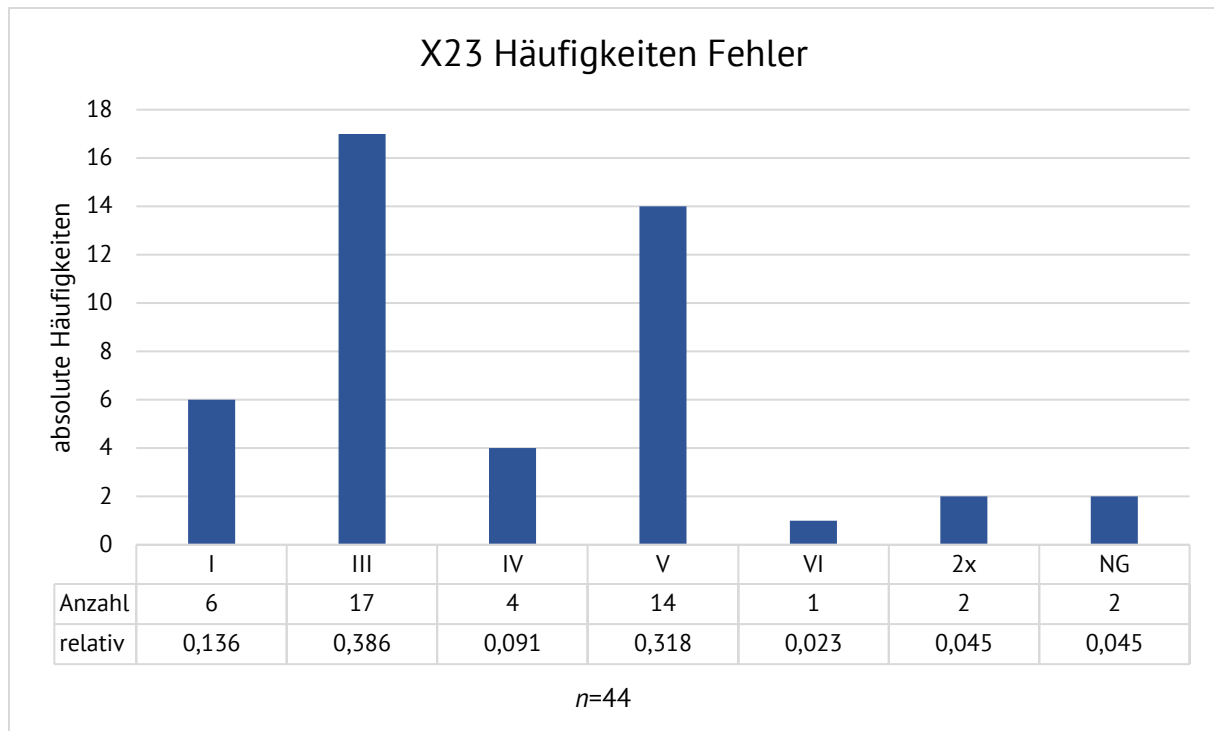
Tabelle 13: X23 Kategoriensystem

4. Auswertung



Statistik 15: X23 Grundgesamtheit: alle Performanzen

Bei 84,7 Prozent korrekt bearbeiteten Aufgaben, sprich 255 Antworten, die alle in dieselbe Kategorie fallen (Statistik 15), steht zu erwarten, dass bei fünf weiteren Antworten und lediglich 46 verbleibenden Performanzen keine der verbleibenden Kategorien besonders stark ausgeprägt sein dürfte. Dazu kommen noch vier weitere Performanzen (2x & NG), die keiner Antwortkategorie zugeordnet werden können. Insofern sind die 17 Entsprechungen für Kategorie III beziehungsweise die 14 Entsprechungen für Kategorie V doch recht eindeutig. Am anderen Ende des Spektrums liegt die Kategorie VI, die lediglich von einer Schülerin oder einem Schüler angekreuzt wurde. Wird die Grundgesamtheit der falsch beantworteten Aufgaben betrachtet (Statistik 16), ist zu sehen, dass 38,6 Prozent der Schülerinnen und Schüler, die die Aufgabe 23 nicht lösen konnten die dritte Antwort gewählt haben. Immer noch 31,8 Prozent haben die fünfte Antwort gewählt



Statistik 16: X23 Grundgesamtheit: fehlerhafte Antworten

Die häufigste Fehlerkategorie (III) entspricht genau den vertauschten Wahrscheinlichkeiten. Hierbei wurde die zu den schwarzen Kugeln gehörige Wahrscheinlichkeit den roten Kugeln zugeordnet und umgekehrt. Bei Kategorie V wurde fälschlicherweise *mindestens eine rote Kugel* mit *genau einer roten Kugel* identifiziert.

5. Besonderheiten

Mit der vorliegenden Aufgabenstellung ließe sich diese Aufgabe auch der Grundkompetenz WS 3.3 zuordnen: „Situationen erkennen und beschreiben können, in denen mit Binomialverteilung modelliert werden kann“ (BIFIE 2013a: 17).

4.12 Aufgabe 24

Vergleich zweier Konfidenzintervalle

Auf der Grundlage einer Zufallsstichprobe der Größe n_1 gibt ein Meinungsforschungsinstitut für den aktuellen Stimmenanteil einer politischen Partei das Konfidenzintervall $[0,23; 0,29]$ an. Das zugehörige Konfidenzniveau (die zugehörige Sicherheit) beträgt γ_1 .

Ein anderes Institut befragt n_2 zufällig ausgewählte Wahlberechtigte und gibt als entsprechendes Konfidenzintervall mit dem Konfidenzniveau (der zugehörigen Sicherheit) γ_2 das Intervall $[0,24; 0,28]$ an. Dabei verwenden beide Institute dieselbe Berechnungsmethode.

Aufgabenstellung:

Ergänzen Sie die Textlücken im folgenden Satz durch Ankreuzen der jeweils richtigen Satzteile so, dass eine korrekte Aussage entsteht!

Unter der Annahme von $n_1 = n_2$ kann man aus den Angaben ① folgern;
unter der Annahme von $\gamma_1 = \gamma_2$ kann man aus den Angaben ② folgern.

①	
$\gamma_1 < \gamma_2$	☐
$\gamma_1 = \gamma_2$	☐
$\gamma_1 > \gamma_2$	☐

②	
$n_1 < n_2$	☐
$n_1 = n_2$	☐
$n_1 > n_2$	☐

1. Grundkompetenz: WS 4.1

Konfidenzintervalle als Schätzung für eine Wahrscheinlichkeit oder einen unbekannten Anteil p interpretieren (frequentistische Deutung) und verwenden können, Berechnungen auf Basis der Binomialverteilung oder einer durch die Normalverteilung approximierten Binomialverteilung durchführen können

2. Beschreibung

Die letzte untersuchte Aufgabe (X24) hat den Vergleich zweier Konfidenzintervalle zum Inhalt. 64 Prozent der Schülerinnen und Schüler konnten diese Aufgabe korrekt lösen. Um den Punkt für diese Aufgabe zu erhalten musste beide Sätze richtig vervollständigt werden (Antwortformat: Lückentext). Dies bedeutet, dass das Kreuz zu Ziffer ① an der dritten Stelle zu setzen ist, während das Kreuz für den zweiten Satz ② an der ersten Stelle zu setzen ist.

3. Kategorien

Die Kategorien ergeben sich jeweils aus den anzukreuzenden Aussagen. Dabei muss zwischen den Aussagen für den ersten Satz ① und den zweiten Satz ② unterschieden werden.

Die Antworten zu ① werden von oben nach unten mit den Buchstaben *i*, *ii* und *iii* bezeichnet und bilden auch die entsprechenden Kategorien (Tabelle 14). Analog dazu werden die drei Antwortmöglichkeiten zu ② mit *I*, *II* und *III* beschriftet (Tabelle 15).

Kategorie	Beschreibung
i	$\gamma_1 < \gamma_2$
ii	$\gamma_1 = \gamma_2$
iii	$\gamma_1 > \gamma_2$

Tabelle 14: X24 Kategoriensystem ①

Kategorie	Beschreibung
I	$n_1 < n_2$
II	$n_1 = n_2$
III	$n_1 > n_2$

Tabelle 15: X24 Kategoriensystem ②

Dies ergibt, ähnlich wie bei Aufgabe 1, verknüpfte Tabellen der möglichen Antwortkombinationen (Tabelle 16 und Tabelle 17).

4. Auswertung

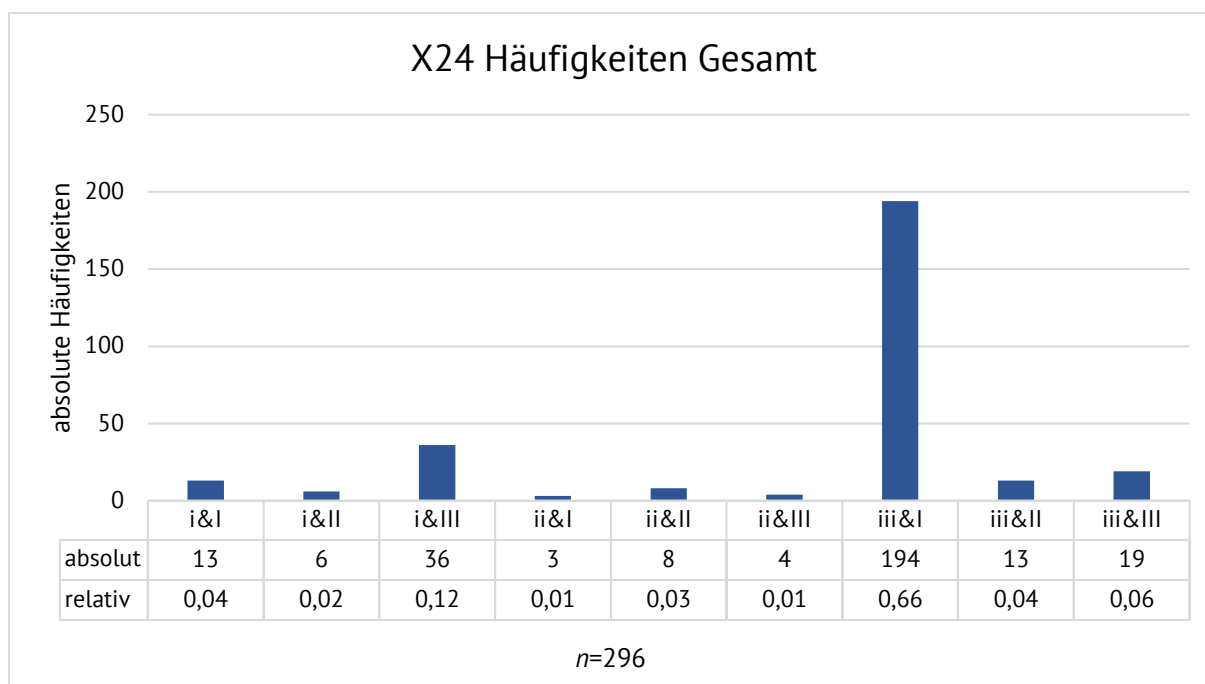
Absolut (n=296)				
&	I	II	III	
i	13	6	36	
ii	3	8	4	
iii	194	13	19	

Tabelle 16: X24 Häufigkeitstabelle (absolut) verknüpfte Antworten (Grundgesamtheit: alle Performanzen)

Relativ (n=296)				
&	I	II	III	
i	0,04	0,02	0,12	
ii	0,01	0,03	0,01	
iii	0,66	0,04	0,06	

Tabelle 17: X24 Häufigkeitstabelle (relativ) verknüpfte Antworten (Grundgesamtheit: alle Performanzen)

Sowohl in den Häufigkeitstabellen (Tabelle 16, Tabelle 17 und Tabelle 18) als auch in Statistik 17 wird mit der Grundgesamtheit $n=296$ gearbeitet. Fünf Schülerinnen und Schüler haben die Aufgabe nicht bearbeitet.



Statistik 17: X24 Grundgesamtheit: alle Performanzen

	①			②		
$n=296$	i	ii	iii	I	I	III
Absolut	55	15	226	210	27	59
Relativ	0,19	0,05	0,76	0,71	0,09	0,20

Tabelle 18: X24 Häufigkeitstabelle einzelne Ankreuzungen (Grundgesamtheit: alle Performanzen)

Die beiden häufigsten falschen Einzelantworten, 19 Prozent der Schülerinnen und Schüler wählten Antwort i für die erste Lücke, 20 Prozent Antwort III für die zweite Lücke, erweisen sich kombiniert auch als die am häufigsten gewählte falsche Antwortkombination. Bei 36 Performanzen (12 Prozent) wurden sowohl i als auch III angekreuzt. Werden diese beiden Aussagen gewählt, kehrt sich der Inhalt, im Vergleich zu den richtigen Antworten, genau um. Sie entsprechen dem Gegenteil der richtigen Antworten.

Anders als bei Aufgabe 1, bei der die Differenz zwischen den richtigen einzelnen Ankreuzungen und der richtigen Kombination recht groß war, weist Aufgabe 24 hier keine großen Unterschiede auf. Der größere Unterschied liegt zwischen der richtigen Einzelantwort iii und der richtigen Kombination. 32 Schülerinnen und Schülerinnen ($226 - 194 = 32$) und somit 11 Prozent (Grundgesamtheit: alle Performanzen) haben beim ersten Satz das richtige Kreuz gesetzt, konnten den zweiten Satz jedoch nicht korrekt beantworten. Umgekehrt

konnten 16 Maturantinnen und Maturanten ($210 - 194 = 16$) den zweiten Satz korrekt vervollständigen, jedoch nicht den ersten.

Schaut man sich nun an in welcher Kombination, in Bezug auf die Grundgesamtheit der nicht korrekt gelösten Aufgaben ($n=102$), die falschen Antworten aufgetreten sind, ergibt sich Tabelle 19:

$n=102$	i&x	ii&x	iii&x	x&I	x&II	x&III
Absolut	55	15	32	16	27	59
Relativ	0,51	0,15	0,31	0,16	0,26	0,58

Tabelle 19: X24 Häufigkeitstabelle Kategorien in falscher Kombination (Grundgesamtheit: fehlerhafte Antworten)

Dabei wurden die richtige Kombination (iii&I) ausgeschlossen. Folglich wählten insgesamt 48 Schülerinnen und Schüler (iii&x+x&I), was einem Anteil von 0,47 entspricht, eine der beiden richtigen Antworten. Diese beiden relativen Anteile dürfen tatsächlich zusammengezählt werden, da sie keine gemeinsamen Kategorien enthalten. Für die anderen kombinierten Kategorien gilt das nicht: So enthalten beispielsweise die beiden kombinierten Kategorien i&x und x&I beide die Kategorie i&I.

Insgesamt wurde Kategorie III, mit 59 Ausprägungen, am öftesten falsch kombiniert. Das heißt, dass 58 Prozent der Schülerinnen und Schüler, die die Aufgabe falsch beantwortet haben, eine Antwort in Verbindung mit Kategorie III gewählt haben. Weiters lässt sich aus Tabelle 19 herauslesen, dass 51 Prozent der Maturantinnen und Maturanten (das sind 55 Schülerinnen und Schüler) ihr Kreuz bei Kategorie i gesetzt haben, was somit der zweithäufigsten kombinierten Kategorie entspricht. Um die Überschneidungen zwischen den Kategorien in Kombination noch weiter zu verdeutlichen: Es gilt zu beachten, dass 36 Performanzen (siehe Tabelle 16) in beide Kategorien in Kombination fallen.

5. Besonderheiten

Durch die Nichtberücksichtigung der nicht gemachten Aufgaben in der Grundgesamtheit ist bei der Auswertung dieser Aufgabe immer von 66 Prozent richtig bearbeiteten Performanzen die Rede. Für die durchschnittliche Lösungsquote werden selbstverständlich auch die nicht gemachten Performanzen hinzugezogen, woraus die in Abbildung 2 korrekt dargestellte durchschnittliche Lösungsquote von 0,64 entsteht.

5 Resümee und Ausblick

Zwölf Aufgaben des Haupttermins der standardisierten schriftlichen Reifeprüfung Mathematik 2016 wurden untersucht und so stehen am Ende zwölf Kategoriensysteme und somit zwölf Ergebnisse. Die deskriptive Darstellung der Kategorien unterscheidet sich jeweils von Aufgabe zu Aufgabe zu Aufgabe, genauso wie auch die Häufigkeitsanalysen jeweils spezifisch sind. Mit diesen zwölf Kategoriensystemen wurde erstmals ein Teil der ssRP Mathematik 2016 mit Augenmerk auf die Performanzen der Schülerinnen und Schüler untersucht.

Es bleibt wichtig festzuhalten, dass die vorgestellten Kategoriensysteme selbstverständlich keine absolute Wahrheit bilden können. Auch hier gilt, neben den Vorteilen von Kategorien, beispielsweise die Zusammenfassung zu übersichtlicheren Clustern, dass Kategorien „auch zu Verzerrungen der Übereinstimmungen und Unterschiede der von ihnen zusammengefassten Objekte führen“ (WALDMANN 2016: 379) können. Eine Kategorie kann von diesem Bedenken allerdings ausgenommen werden: Die im Bedarfsfall geführte *Kategorie Nicht Gemacht* bildet trennscharf genau die nicht bearbeiteten Aufgaben ab. Damit kann zwischen den bearbeiteten, aber falsch bearbeiteten und den nicht gemachten Aufgaben unterschieden werden. Nur ein Beispiel für den Mehrwert der erarbeiteten Kategorien. Eine solche Unterscheidung war am Anfang dieser Arbeit, also aus den Rohdaten, nicht möglich.

Mit diesem Gewinn an Informationen kann nun vor allem eines passieren: weiter damit zu arbeiten. Die Untersuchung der Antworten ermöglicht tiefergehende Analysen, beispielsweise zu den Aufgabenstellungen oder auch zu den Formaten der Aufgaben. Auch vergleichende Analysen zu anderen Terminen oder anderen Schulformen werden möglich. Mithilfe von einschlägiger Fachliteratur könnten auch allgemeingültige Fehlerkategorien, unabhängig von spezifischen Aufgaben, gebildet werden.

Demselben Ansatz folgend, aber mit anderem Umfang, können Grundlagen für Vergleiche oder auch für den Anspruch der vollständigen Analyse einer standardisierten schriftlichen Reifeprüfung Mathematik geschaffen werden. So können beispielsweise die Aufgaben des Teil 1, die hier noch offen geblieben sind (Aufgaben 7-18) nach dem Vorbild dieser Arbeit ebenso kategorisiert werden. Auch eine Kategorisierung der Performanzen des Teil 2 ist möglich. Genauso kann der Umfang auch auf andere Termine (Nebentermine), Jahre,

Schulformen (zum Beispiel: BORG) oder Schultypen (BHS) ausgeweitet werden. Insbesondere ein Vergleich mit Fehlerkategorien ab dem Haupttermin 2018, ab dem eine Minimalanforderung für elektronische Hilfsmittel verlangt wird, wäre interessant.

Möglichkeiten zu weiteren Arbeiten in diesem Themengebiet gäbe es demnach genügend und durch die Erhebung der Daten seitens des Bundesministeriums für Bildung, Wissenschaft und Forschung auch die Chance dazu zu arbeiten. So besteht die Gelegenheit, die standardisierte schriftliche Reifeprüfung Mathematik wissenschaftlich fundiert weiterzuentwickeln, wozu auch diese Arbeit einen kleinen Teil beigetragen hat. Mit weiterführender Forschung zu den Ergebnissen dieser Diplomarbeit kann dieser Teil auch noch ausgebaut werden.

Literaturverzeichnis

BAUR, N. & BLASIUS, J. (Hg.) (2014a): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS.

BAUR, N. & BLASIUS, J. (2014b): Methoden der empirischen Sozialforschung. Ein Überblick. In: N. BAUR & J. BLASIUS (Hg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, S. 41-62.

BIFIE (2013a): *Die standardisierte schriftliche Reifeprüfung Mathematik. Inhaltliche und organisatorische Grundlagen zur Sicherung mathematischer Grundkompetenzen*. AUE, V.; FREBORT, M.; HOHENWARTER, M.; LIEBSCHER, M.; SATTLBERGER, E.; SCHIRMER, I.; SILLER, H.-S. (Leitung); VORMAYR, G.; WEIß, M.; WILLAU, E. Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. Online verfügbar: https://www.srdp.at/fileadmin/user_upload/downloads/Begleitmaterial/07_MAT/srdp_ma_konzept_2013-03-11.pdf (02.02.2018).

BIFIE (2013b): *Praxishandbuch Mathematik AHS Oberstufe. Auf dem Weg zur standardisierten kompetenzorientierten Reifeprüfung. Teil 1. 2. aktualisierte Auflage*. BREYER, G.; FÜRST, S.; HEUGL, H.; KRAKER, M.; LIEBSCHER, M. (Projektleiterin); PREIS, C.; SVECNIK, E.; LIEGL, I.; PLATTNER, G. Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. Online verfügbar: https://www.srdp.at/fileadmin/user_upload/downloads/Begleitmaterial/07_MAT/Publikationen/srdp_ma_praxishandbuch_mathematik_teil_1.pdf (07.02.2018).

BIFIE (2016a): *Mathematik Teil-1-Aufgaben. Standardisierte kompetenzorientierte schriftliche Reifeprüfung AHS*. Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. Online verfügbar: https://www.srdp.at/fileadmin/user_upload/downloads/Matura_2015-16/12_MAT/KL16_PT1_AHS_MAT_T1_CC_AU.pdf (04.05.2018).

BIFIE (2016b): *Mathematik Teil-1-Aufgaben Korrekturheft. Standardisierte kompetenzorientierte schriftliche Reifeprüfung AHS*. Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens. Online verfügbar: https://www.srdp.at/fileadmin/user_upload/downloads/Matura_2015-16/12_MAT/KL16_PT1_AHS_MAT_T1_CC_LO.pdf (04.05.2018).

BLASIUS, J. (2014): Skalierungsverfahren. In: N. BAUR & J. BLASIUS (Hg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, S. 1051-1062.

BRÜSEMEISTER, T. (2008): *Qualitative Forschung. Ein Überblick*. Wiesbaden: VS Verlag für Sozialwissenschaften.

DANGL, M., FISCHER, R., HEUGL, H., KRÖPFL, B., LIEBSCHER, M., PESCHEK, W. & SILLER, H.-S. (2009): *Das Projekt "Standardisierte schriftliche Reifeprüfung aus Mathematik". Sicherung von mathematischen Grundkompetenzen*. Österr. Kompetenzzentrum für Mathematikdidaktik. Online verfügbar: <https://www.aau.at/didaktik-der-mathematik/publikationen/bildungsstandards-zentralmatura/materialien-berichte/#MaterialienZM> (07.02.2018).

ERNST, P. (2011): *Germanistische Sprachwissenschaft. Eine Einführung in die synchrone Sprachwissenschaft des Deutschen*. Wien: Facultas.wuv.

FLICK, U. (1995): Stationen des qualitativen Forschungsprozesses. In: U. FLICK, E.V. KARDORFF & H. KEUPP (Hg.): *Handbuch qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen*. Weinheim: Beltz, S. 147-173.

GLASER, B.G. & STRAUSS, A.L. (2010): *Grounded theory. Strategien qualitativer Forschung*. Bern: Verlag Hans Huber.

HÄDER, M. & HÄDER, S. (2014): Stichprobeziehung in der quantitativen Sozialforschung. In: N. BAUR & J. BLASIUS (Hg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, S. 283-297.

HARTMANN, F. (2008): *Medien und Kommunikation*. Wien: Facultas.

HOLLAND, H. & SCHARNBACHER, K. (2010): *Grundlagen der Statistik. Datenerfassung und -darstellung, Maßzahlen, Indexzahlen, Zeitreihenanalyse*. Wiesbaden: Gabler Verlag / Springer Fachmedien.

HUG, T. & POSCHESCHNIK, G. (2015): *Empirisch forschen. Die Planung und Umsetzung von Projekten im Studium*. Konstanz: UVK.

KAYA, M. & HIMME, A. (2009): Möglichkeiten der Stichprobenbildung. In: S. ALBERS, D. KLAPPER, U. KONRADT, A. WALTER & J. WOLF (Hg.): *Methodik der empirischen Forschung*. Wiesbaden: Gabler Verlag, S. 79-88.

KELLE, U. (2008): *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage.

KLIX, F. (1995): Stabilität und Wandlungen in geistigen Dispositionen des Menschen. *Sitzungsberichte der Leibniz-Sozietät*, S. 5-40.

KRAMER, S. & SATTLBERGER, E. (2016): Habuemus Haupttermin. Die SRP in Mathematik (AHS) nach 2015 - wie war's und wie geht's weiter? In: ÖSTERREICHISCHE MATHEMATISCHE GESELLSCHAFT (Hg.): *Schriftenreihe zur Didaktik der Mathematik der Österreichischen Mathematischen Gesellschaft. Vorträge der 38. Lehrerfortbildung am 1. April 2016 an der Fakultät für Mathematik der Universität Wien*. Wien: Stadtschulrat für Wien, S. 60-73. Online verfügbar: <https://www.oemg.ac.at/DK/Didaktikhefte/2016%20Band%2049/VortragKramerSattlberger.pdf> (09.07.2018).

MAIER, H. & SCHWEIGER, F. (1999): *Mathematik und Sprache. Zum Verstehen und Verwenden von Fachsprache im Mathematikunterricht*. Wien: öbv & hpt.

MAYRING, P. (2015): *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz.

MAYRING, P. (2016): *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. Weinheim, Basel: Beltz.

MIKA, T. & STEGMANN, M. (2014): Längsschnittanalyse. In: N. BAUR & J. BLASIUS (Hg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS, S. 1077-1087.

PESCHEK, W. (2011): Sicherung mathematischer Grundkompetenzen am Beispiel des österreichischen Zentralabiturs. In: M.A. HELMERICH, K. LENGNINK, G. NICKEL & M. RATHGEB (Hg.): *Mathematik verstehen. Philosophische und didaktische Perspektiven*. Wiesbaden: Vieweg+Teubner Verlag / Springer Fachmedien, S. 211-220.

REINHOFFER, B. (2008): Lehrkräfte geben Auskunft über ihren Unterricht. Ein systematisierender Vorschlag zur deduktiven und induktiven Kategorienbildung in der Unterrichtsforschung. In: P. MAYRING & M. GLÄSER-ZIKUDA (Hg.): *Die Praxis der qualitativen Inhaltsanalyse*. Weinheim: Beltz, S. 123-141.

RIESENHUBER, F. (2009): Großzahlige empirische Forschung. In: S. ALBERS, D. KLAPPER, U. KONRADT, A. WALTER & J. WOLF (Hg.): *Methodik der empirischen Forschung*. Wiesbaden: Gabler Verlag, S. 1-16.

ROSCH, E. (1978): Principles of Categorization. In: E. ROSCH & B.B. LLOYD (Hg.): *Cognition and categorization. Conference*. Hillsdale, NJ: Lawrence Erlbaum Associates, S. 27-48.

TIEMANN, V. (2012): *Statistik. Für Studienanfänger*. Stuttgart: UTB.

WALDMANN, M. R. (2016): Kategorisierung und Wissenserwerb. In: J. MÜSSELER (Hg.): *Allgemeine Psychologie*. Berlin, Heidelberg: Springer, S. 377-427.

Abstract

Die vorliegende Diplomarbeit setzt sich mit dem Haupttermin der standardisierten schriftlichen Reifeprüfung Mathematik 2016 in Österreich auseinander. Dabei wird mit dem Kriterium der Lösungsquote eine Auswahl an Aufgaben getroffen, die analysiert wird. Das Augenmerk wird auf die Ausarbeitungen der Schülerinnen und Schüler gelegt, welche mithilfe der induktiven Kategorienbildung in Bezug auf Fehler kategorisiert werden. Die so entstandenen Kategorien werden durch eine Häufigkeitsanalyse weiter ausgewertet. Somit stehen am Ende der Arbeit zwölf Kategoriensysteme, welche sowohl die große Bandbreite an Bearbeitungsmöglichkeiten in manchen Fällen als auch die uniforme Herangehensweise in anderen Fällen aufzeigen.

This thesis deals with the written mathematics exam of the centralized A-levels in Austria in 2016. The solution quote selects a set of tasks that will be analysed. The focus herein lies on the students' elaborations. By means of the method of inductive categorization those elaborations get categorized in relations to errors. The resulting categories are further evaluated by a frequency analysis. The result of the thesis are twelve category systems, which show both: The wide range of different approaches on one task in some cases and the uniform strategies in other cases.