# Bedingungsfaktoren für den erfolgreichen Übergang von Schule zu Hochschule

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen
Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen

im Promotionsstudiengang „Mathematical Sciences“
der Georg-August University School of Science (GAUSS)

vorgelegt von
**Kolja Pustelnik**
aus Leer

Göttingen, 17. August 2018

Betreuungsausschuss

Prof. Dr. Stefan Halverscheid, Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Susanne Schneider, Didaktik der Physik, Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Stefan Halverscheid, Mathematisches Institut, Georg-August-Universität Göttingen

Korreferentin: Prof. Dr. Susanne Schneider, Mathematisches Institu, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Ina Kersten, Mathematisches Institut, Georg-August-Universität Göttingen

Jun.-Prof. Dr. Christoph Lehrenfeld, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen

Prof. Dr. Walther Paravicini, Mathematisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Anja Sturm, Institut für Mathematische Stochastik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 28.09.2018

# Inhaltsverzeichnis

# Tabellenverzeichnis

# 1 Einleitung

Der Studienabbruch in mathematisch-naturwissenschaftlichen Fächern stellt an deutschen Universitäten weiterhin ein bedeutsames Phänomen dar. Dazu stellen Heublein et al. (2012) und Dieter (2012) Zahlen der Abbrecherquoten dar. Dabei stellt bereits die exakte Bezifferung des Abbruchs eine erste Schwierigkeit dar und wird entsprechend von beiden Studien unterschiedlich gelöst, sodass sich auch sehr unterschiedliche Zahlen ergeben.
In Heublein et al. (2012) wird die Studienabbruchquote eines Fachs definiert als der Anteil der Studierenden, die ihr erstes Studium in diesem Fach aufgenommen haben und keinen Abschluss erreicht haben. Das Fach des Abschlusses nicht berücksichtigt. Im Vergleich der Fächergruppen ergibt sich für Mathematik und Naturwissenschaften mit 28 % die höchste Abbruchquote. Dabei zeichnet sich diese Fächergruppe darüber hinaus dadurch aus, dass die als einzige einen Anstieg der Abbruchquote im Vergleich der letzten sieben Jahre besitzt. Innerhalb der Fächergruppe besitzen wiederum Mathematik, Physik, Chemie und Informatik mit jeweils über 30 % Abbrecherquote den höchsten Anteil an Abbrechern.
Da diese Quoten nicht berücksichtigen, in welchem Fach der Abschluss erzielt wurde, werden die Verluste an Studierende im Bezug auf die Fächergruppe nicht aufgenommen. Dazu werden in Dieter (2012) bezogen auf Mathematik die Studienabbruchquoten durch eine Saldenrechnung errechnet. Die Abbrechenden ergeben sich dann als die Differenz aus Immatrikulationen zu einem Fachsemester und den Immatrikulierten aus dem vorigen Fachsemester im vorigen Semester. In der Berechnung werden frühe und späte Studienabbrechenden unterschieden. Für die frühen, das heißt bis zum Beginn des dritten Semesters ergibt sich für das letzte untersuchte Anfangsjahr 2005 eine Quote von 30,9 % und für die Studienabbruchquote nach acht Fachsemestern für das Anfangsjahr 2005 eine Quote von 70,1 % für Frauen und von 59,7 % für Männer. Von Jahr zu Jahr ergeben sich dabei nach dem dritten Fachsemester immer Abbruchquoten, die über 10 % leigen. Diese Zahlen zeigen aus Sicht des Studiengangs Mathematik ein deutlich schlechteres Bild, wenn über 60 % der Studienanfängerinnen und -anfänger keinen Masterabschluss erreichen.
Es zeigt sich also, dass der Studienabbruch in Mathematik sich nicht nur auf die Studieneingangsphase beschränkt, sondern ein Phänomen der gesamten Studiendauer darstellt. Dennoch stellt das erste Jahr mit der deutlich höchsten Abbruchquote den zentralen Einzelaspekt dar. Daher wird sich hier auf den Abschnitt des ersten Semesters konzentriert.
Das Phänomen des Studienabbruchs stellt natürlich nicht ein ausschließlich deutsches Phänomen dar, sondern findet auch international Beachtung. Hierbei zeigen sich wesentliche Unterschiede zwischen den verschiedenen Staaten und den zugehörigen Schul- und Studiensystemen. Dies führt auch zu deutlichen Unterschieden in Ursachen und möglichen Lösungsansätzen. Dement-

sprechend konzentriert sich diese Arbeit auf eine deutsche Perspektive in bisherigen Ergebnissen.
Es ist der Studienabbruch auf der einen Seite mit hohen Kosten für die Gesellschaft verbunden, aber auch für den oder die Abbrechende. Daher wurden in den letzten Jahren in Deutschland verschiedene Maßnahmen ergriffen, insbesondere den Studieneingang in mathematikhaltigen Studiengängen zu erleichtern. Eine Übersicht über verschiedene Maßnahmen und Ansätze geben zum Beispiel (Bausch et al., 2014; Hoppenbrock et al., 2016). Es zeigt sich, dass das Abbruchphänomen für verschiedene Studiengänge und Institutionen, wie global Universität im Vergleich mit Fachhochschulen, sich deutlich in ihren Gründen unterscheiden.
In dieser Arbeit soll entsprechend ein sehr fokussierter Blick auf wenige Studiengänge genommen werden. Das bedeutet, dass nur wenige Studiengänge untersucht werden, welche einen ähnlichen mathematischen Hintergrund besitzen. Dabei konzentriert sich die Untersuchung auf die Studiengänge "Mathematik" , "Physik", "angewandte Informatik" sowie Lehramt Mathematik. Denn diese Studiengänge besitzen den höchsten Anteil an Mathematik im Vergleich mit anderen Studiengängen und besuchen daher einen gemeinsamen mathematischen Vorkurs, in welchem auch die Daten erhoben werden. Bezüglich der mathematischen Fähigkeiten lässt sich also eine relativ große Ähnlichkeit im Vergleich zu anderen Studiengängen erwarten.
Neben den individuellen Faktoren spielen natürlich auch institutionelle Faktoren eine wesentliche Rolle für das Studium und dessen Eingangsphase. Diese Variablen werden in dieser Studie allerdings nicht weiter untersucht, sodass sich die Untersuchung nur auf eine Universität konzentriert, und so Varianz in institutionellen Hintergründen aus den Analysen heraushält. Dies stellt auf der einen Seite natürlich eine Einschränkung dar, erlaubt auf der anderen Seite aber auch das Herausarbeiten spezifischer Zusammenhänge an einem Standort.
Von den Studienabbrecherinnen und -abbrechern selbst werden leistungsbezogene Gründe als die Hauptursachen des Fachwechsels oder Abbruchs angegeben. Auch diese Angaben machen die Fachspezifität des Abbruchs deutlich, da der jeweilige Studiengang die Leistungsansprüche an die Studierenden vorgibt. Es stellt sich nun die Frage, in wie fern die Eingangsvoraussetzungen der Studierenden für diese fachspezifischen Probleme eine Bedeutung besitzen bzw. wie stark diese neu an der Universität entstehen. Insbesondere ist auch interessant, wie stark gebietsabhängig sich die Eingangsvoraussetzungen der Studierenden unterscheiden und, ob spezifische mathematische Bereiche identifizierbar sind, die besonderen Einfluss besitzen bzw. für die sich besondere Unterschiede messen lassen.
Dementsprechend stellen die fachmathematischen Eingangsvoraussetzungen einen zentralen Aspekt des Übergangs dar. Die Eingangsvoraussetzungen werden durch einen dafür entwickelten Test erhoben. Die Eingangsvoraussetzungen werden charakterisiert als in der Schule zu erwerbendes Wissen,

welches für die Universität Relevanz besitzt. Die Eingangsvoraussetzungen der Studienanfängerinnen und -anfänger sollen dabei daraufhin untersucht werden, von welchen individuellen Faktoren sie abhängig sind.
Neben den Eingangsvoraussetzungen der Studierenden stellt sich die Frage, wie sich die Studierenden in Abhängigkeit davon am Studienbeginn entwickeln, der hier fokussiert wird. Es stellt sich die Frage, was ein erfolgreiches Studium ausmacht. In den beiden oben beschriebenen Studien wurde das Abbruchsverhalten durch Immatrikulation definiert. Dies beschreibt allerdings nicht das Vorankommen der Studierenden innerhalb ihres Studiums. Daher werden hier als Erfolgskriterium wiederum leistungsbezogene Variablen, also die Klausuren des ersten Semesters betrachtet werden. Dazu gehören das Antreten zu einer Klausur, das Bestehen der Klausur sowie die Zensur in der Klausur, welche jeweils für die beiden Veranstaltungen des ersten Semesters, "Differential- und Integralrechnung" und "Analytische Geometrie und lineare Algebra" betrachtet werden.
Damit betrachtet die Arbeit in einem fokussierten Blick für wenige Studiengänge spezifisch die leistungsbezogenen Eingangsvoraussetzungen sowie ebenso leistungsbezogen den Erfolg im ersten Semester des Studiums. Eine Arbeit, welche sich mit einem breiteren Blick auf verschiedene mathematikhaltige Studiengänge bezieht und auch weitere individuelle Persönlichkeitsmerkmale in die Analysen aufnimmt.
In dieser Arbeit wird zunächst ein theoretisches Modell für die Betrachtung des Studieneinstiegs als Übergangsritus vorgestellt. Dieses geht von einem Bruch, und nicht einem Übergang, zwischen Schule und Universität aus. Dieses wird explizit in Bezug auf die Mathematik sowie das Lernen an der Universität anhand einer Beschreibung der Schwierigkeiten, welche den Studierenden zum Studienbeginn begegnen.
Anschließend werden bisherige Ergebnisse zu Unterschieden in den mathematischen Fähigkeiten von Schülerinnen und Schüler bzw. Studienanfängerinnen und -anfängern in Abhängigkeit von den erhobenen Variablen beschrieben. Diese Darstellung wird sich wie oben beschrieben so weit möglich auf die Ergebnisse deutscher Studien beschränkt. Hierbei werden explizit die acht verschiedenen mathematischen Gebiete dargestellt.
In Abschnitt 5 wird die Entwicklung des Testinstruments der mathematischen Eingangsvoraussetzungen der Studienanfängerinnen und -anfänger vorgestellt. Dieses stellt in seiner Konzeption einen Überschnitt von schulischem und universitären Wissen dar. Insbesondere wird eine Einordnung der verschiedenen mathematischen Gebiete dargestellt.
Als Stichprobe der Untersuchung dienen die Teilnehmenden des Mathematische Propädeutikums. Am ersten Tag der Veranstaltung bearbeiten die Teilnehmenden den neunzigminütigen Test. Die Stichprobe besteht aus den Teilnehmenden der Jahre 2013 bis 2016. Als Vergleichsgruppe wurden weiterhin Schülerinnen und Schüler aus der Umgebung untersucht. Die Auswertung in den nächsten beiden Abschnitten geht dann der Frage nach, wovon

die Eingangsvoraussetzungen abhängen, und wie sich diese auf die Klausurergebnisse im ersten Semester auswirken. Dabei werden zunächst die Unterschiede in den Eingangsvoraussetzungen in Abhängigkeit einzelner Variablen untersucht und anschließend ihr gemeinsamer Einfluss. Ein entsprechendes Vorgehen wird für das Bestehen der Klausuren im ersten Semester verfolgt. Hier wird längsschnittlich untersucht, welchen Einfluss die vor dem Studium erhobenen Variablen auf den Klausurerfolg besitzen, bzw. wie gut der Klausurerfolg vorhergesagt werden kann.

Für beide Teile der Auswertung, die Voraussetzungen sowie die Klausurergebnisse, werden dabei insbesondere domänenspezifische Auswertungen durchgeführt. Dies bedeutet, dass Unterschiede in den acht Bereichen des mathematischen Tests herausgearbeitet werden, insbesondere zeitliche Effekte spielen hier eine Rolle. Ebenso werden die beiden Veranstaltungen des ersten Semesters getrennt ausgewertet und verglichen.

In den letzten beiden Abschnitten werden die Ergebnisse der Auswertungen diskutiert und die Forschungsfragen erörtert. Anschließend werden die Einschränkungen der Studie sowie mögliche Ausblicke gegeben.

# 2 Theoretischer Hintergrund

Der Übergang zwischen der Schule und der Universität wurde bereits aus verschiedenen Perspektiven betrachtet. Für die jeweiligen Untersuchungen wird dabei in der Regel ein jeweiliger theoretischer Hintergrund gewählt. Einen Überblick über einige der Forschungsrichtungen gibt (Gueudet, 2008), der insbesondere den Übergang in persönliche, institutionelle und soziale Phänomene unterteilt. Entsprechend werden in diesen Untersuchungen einzelne wichtige Phänomene des Übergangs in den Fokus genommen. Einen globalen Rahmen für die Einordnung des Übergangs für die Mathematik aus anthropologischer Perspektive wird dann in (Clark & Lovric, 2009) gegeben. Zentrales Merkmal ist dabei das Ausgehen eines klaren Bruchs zwischen der Schule und der Universität im Gegensatz zu einem graduellen Übergang. Anschließend sollen dann, im Sinne des Fokusses dieser Arbeit auf die mathematischen Fähigkeiten, diese vor diesem Hintergrund beschrieben werden. Dabei werden die Änderungen der Mathematik an der Universität beschrieben auf Grundlage von verschiedenen Untersuchungen zu Schwierigkeiten, die Studienanfängerinnen und -anfänger aufweisen.

## 2.1 Übergang als Übergangsritus

In (Clark & Lovric, 2009) und weiter entwickelt in (Clark & Lovric, 2008) beschreiben die Autoren den Übergang zwischen Schule und Universität als "rite of passage" bzw. Übergangsritus. Dieses Modell geht auf Van Gennep (2013) zurück und beschreibt allgemein Lebenssituationen, in welchen eine Krise stattfindet, durch welche das Leben eines Menschen grundsätzlich verändert wird. Grundlegend für den Übergangsritus ist dabei die Feststellung, dass eine Gesellschaft in verschiedene Subgruppen zerfällt, in denen ein Mensch Mitglied ist. Der Übergang lässt sich dann in drei Phasen einteilen. Die erste Phase stellt die Separation da, welche die Trennung von der bisherigen Gemeinschaft beschreibt. Im Bezug auf den Übergang zwischen Schule und Universiät bedeutet dies also die Trennung von der bekannten Schule. Diese findet noch in der Schulzeit statt und stellt die Vorbereitung auf das zukünftige Betreten der Universität dar. Die zweite Phase beschreibt den Übergang selbst und wird liminale Phase genannt. In dieser Phase finden die Veränderungen bzw. die Anpassung an die Universität der Studienanfängerinnen und -anfänger statt. So werden die Routinen verändert und die Person lernt die neue Gemeinschaft an der Hochschule mit ihren veränderten Gewohnheiten, Erwartungen und Regeln kennen. Dabei eignet sich die Person diese, unter Unterstützung der neuen Gemeinschaft, an. Die Integration als letzte Phase findet dann im Laufe des ersten Jahres an der Universität statt, wenn die Studierenden ein Teil der Gemeinschaft an der Universität mit ihren Traditionen, Erwartungen und Regeln geworden sind. Ist die letzte Phase vollzogen worden, besteht wieder eine Balance zwischen der neuen

Gemeinschaft und dem Individuum selbst. Dabei ist entscheidend, dass der vorgenommene Wechsel zwischen zwei klar definierten Zuständen geschieht. Eine Folgerung aus dieser Beschreibung ist, dass der Übergang notwendig als ein Bruch zwischen zwei Gemeinschaften gesehen wird. Insbesondere bedeutet dies, dass der Übergang nicht als ein glatter Übergang verstanden werden kann. Stattdessen ist dies auch nicht wünschenswert, sondern der Übergang stellt einen "Schock" dar, mit dem die Studienanfängerinnen und -anfänger umgehen sollen und müssen. Dementsprechend schlagen die Autoren vor, dass auch Anstrengungen, welche versuchen wollen, den Übergang möglichst sanft zu gestalten, zum Beispiel im Bezug auf die mathematische Sprache, nicht hilfreich sind. Dies gilt ebenso andersherum: Das Simulieren von Universitätsmathematik an der Schule wird als nicht hilfreich gesehen. Es ist also nicht in allen Aspekten hilfreich die Praktiken von Schule und Universität anzunähern.
Dies bedeutet dennoch nicht, dass der Übergang unvorbereitet geschehen muss. Ganz im Gegenteil kommt der Vorbereitung als Teil der Ablösung eine wichtige Funktion zu. Hierzu zählen, neben Informationen zur Vorbereitung auf die Universität, die fachmathematischen Vorbereitungen, welche an der Schule stattfinden. Es müssen allerdings beide Situationen, an der Universität und an der Schule, klar definiert sein.
Als zweite Konsequenz ergibt sich, dass der Übergang an die Universität Zeit braucht. Die Übergangsphase wird hier über einen Zeitraum eines ganzen Jahres gesehen. Dabei ist es eine andauernde Aufgabe diesen Übergang zu gestalten und kann nicht beschleunigt werden. Dies wird insbesondere für das Aufrufen von Inhalten der Schule, die einige Zeit vorauslagen, bzw. für die Zeit, die zwischen dem Schulbesuch und dem Universitätsbesuch liegt, formuliert.
Die dritte Konsequenz erwartet die Selbstständigkeit der Studierenden, denen kein zu großer Teil der Eigenverantwortung abgenommen werden soll durch die Universität, auch wenn dies ungewohnt und weniger komfortabel scheint.
Als vierte Konsequenz wird das symbolische Sterben und Wiederaufstehen gesehen. So gibt es Verhaltensweisen oder auch Wissen aus der Schule, welches für die Universität aufgegeben werden müssen. Beispiele sind ein oberflächliches Lernen oder auch die Definition einer Tangente als Linie, welche den Graph nur einmal berührt. Hierzu gehört insbesondere auch das mathematische Arbeiten sowie die mathematische Sprache, wie sie an der Universität verwendet werden.
In (Clark & Lovric, 2008) wird dieses Modell aufgenommen und in weiteren Aspekten die einzelnen Aspekte von oben insbesondere mit Handlungsempfehlungen versehen. Diese sollen hier nicht weiter ausgeführt werden. Bezogen auf den Bruch in der Mathematik selbst, der zwischen Schule und Universität liegt, ist insbesondere die vierte Konsequenz zentral. Dies deutet bereits das gegebene Beispiel an. Ebenso liegt hier gerade der Fokus dieser Arbeit,

welche auf der einen Seite das Vorwissen der Studienanfängerinnen und -anfänger erhebt und auf der anderen Seite dessen Prädiktion für Erfolg im ersten Semester verwendet. Dieser Fokus auf die Änderung der Mathematik zwischen Schule und Universität wird nun im nächsten Abschnitt anhand zweier Übersichtsartikel gelegt.

## 2.2 Rolle der Mathematik im Übergang zur Universität

Untersuchungen zu Schwierigkeiten im Übergang von Studierenden zur Universität wurden von Gueudet, wie oben erwähnt, zusammengefasst (Gueudet, 2008). Eine entsprechende Übersicht wurde dann vor drei Jahren noch einmal zusammengestellt (Thomas et al., 2015). Entlang ausgewählter Untersuchungen dieser beider Übersichten sollen nun Unterschiede von Mathematik an der Schule und an der Universität herausgearbeitet werden.
Gueudet teilt in ihrer Übersicht die Schwierigkeiten des Übergangs in drei Blöcke ein: „Thinking modes and knowledge's organization", „Proofs and mathematical communication" und „Didactical transposition and didactical contract ". Zu den ersten beiden Abschnitten sollen hier nur in kurzen Ausschnitten zentrale Aspekte wiedergegeben werden. Der instituionelle Rahmen soll dabei nicht berücksichtigt werden, da in dieser Arbeit der Fokus auf einer Universität liegt und die Rahmenbedingungen nicht untersucht werden.
Der erste Abschnitt beschreibt dabei insbesondere die geänderten Anforderungen an Mathematik an der Universität und geht auf Tall (1991) zurück. Es geht in diesem Sinne um den Wechsel von beschreibender und plausibel machender Mathematik hin zu einer Mathematik mit formalen Definitionen und Beweisen. Mathematik ist an der Universität also durch Deduktion und formale Strenge gekennzeichnet. Eine ähnliche Unterscheidung zwischen zwei Arten der Mathematik nehmen auch durch Sierpinska (2000) und Lithner (2003) vor. Dabei geht es auf der einen Seite um eine theoriebezogene und strenge Form des Denkens und auf der anderen Seite eine auf Erfahrungen und Beispielen beruhende Art des Denkens. Während erfahrene Mathematikerinnen und Mathematiker beide Arten flexibel einsetzen können, bleiben Studienanfängerinnen und -anfänger bei der zweiten Form des Denkens, was ihnen an der Universität Schwierigkeiten bereitet. Sierpinska zeigt dieses am Beispiel der lineare Abbildungen.
Einen weiteren klaren Bruch zu Beginn des Studiums stellt die mathematische Sprache und insbesondere die Verwendung der Quantoren dar. Das Erlernen der neuen Sprache fokussiert sich dabei zu Beginn oft auf die Übernahme der oberflächlichen Eigenschaften und weniger auf den Inhalt (Iannone & Nardi, 2007). Der mathematische Formalismus kann dabei als das Lernen einer neuen Sprache gesehen werden (Berger, 2004).
Zusammenfassend zeigen sich also in den betrachteten Untersuchungen, dass es klare Brüche im mathematischen Denken, der Sprache und insbesondere für die Rolle des Beweisens gibt. Dabei zeigen sich diese in ganz verschiede-

nen Situationen und mit verschiedenen Folgen. Die Zusammenstellung von (Thomas et al., 2015) ordnet entsprechend die vorgestellten Untersuchungen nach den mathematischen Fachgebieten sowie in Beweisen und Anwendung. Als mathematische Bereiche werden Analysis, abstrakte Algebra sowie lineare Algebra betrachtet, die auch in Deutschland die Inhalte des ersten Semesters wesentlich abdecken. Für die Analysis zeigt sich, dass die lokale Sichtweise, welches für zum Beispiel Stetigkeit und Ableitung zentral sind, die punktweise Zuordnung und die globale Sichtweise auf Funktionen ablöst (Vandebrouck, 2011). Der bereits in der Schule zentrale Funktionenbegriff muss also umgedeutet werden. Für den ebenso zentralen Folgenbegriff greifen viele Studierende auf Metaphern zurück, welche zum Beispiele physikalische Begründungen oder Unendlich als Zahl verwenden (Oehrtman, 2009). Die Brüche bezogen auf die Analysis zwischen Schule und Universität stellen auch (Bloch & Ghedamsi, 2004) in neun Variablen dar. Diese bestehen unter anderem in Formalisierung, Validierung und Generalisierung, aber auch in der Art von Aufgaben. Es zeigen sich also hier in dem zentral in der Schule vertretenen Gebiet der Analysis klare Brüche bzw. Neudeutungen bekannter Begriffe.
Für die abstrakte Algebra ergeben sich ebenso für viele Begriffe aus der Schule neue Deutung zu Beginn des Studiums: Dies gilt zum Beispiel für das Gleichheitzeichen, die Äquivalenzrelation und Weitere, wie Stadler (2011) zeigt. Hinzu kommt die völlig neue Form der Definitionen algebraischer Strukturen alleine über ihre Eigenschaften, wie im Fall von Gruppen und Körpern, für die wiederum Sprache eine zentrale Bedeutung besitzt (Nardi, 2007, 2011). Im Bezug auf die lineare Algebra zeigten (Stewart & Thomas, 2009, 2010) für zentrale Begriffe wie Lineare Unabhängigkeit und Eigenwerte eine sehr symbolische und prozessorientierte Sicht. Im Vergleich zur Analysis zeigen sich also auch hier in allen auch aus der Schule bekannten Bereichen wesentliche Brüche, dazu kommen aber zusätzlich abstrakte Begriffsbestimmungen.
Es zeigen sich also über alle Gebiete hinweg spezifische Übergangsschwierigkeiten von Studienanfängerinnen und -anfängern. Diese lassen sich im Bezug auf Sprache und mathematische Strenge von Beweisen allgemein fassen. Dabei steht insbesondere die mathematische Deduktion und Axiomatisierung im Vordergrund der Universitätsmathematik gegenüber der Schulmathematik, welche oft auf spezifische Beispiele setzt und plausibele Erklärungen zulässt. Ebenso zeigen sich die Unterschiede aber auch spezifisch in einzelnen Begriffen bzw. Teilgebieten. Dabei sind auch hier Begriffe zu unterscheiden, welche bereits in der Schule behandelt worden sind, aber einer neuen Konzeptionalisierung bedürfen, wie Funktionen, und auf der anderen Seite völlig neue Begrifflichkeiten wie Gruppen und Körper. Zusammenfassend zeigt sich also, dass sich die Schulmathematik und die Universitätsmathematik in ihrem Charakter wesentlich unterscheiden. Dabei bestehen nicht nur graduelle Unterschiede, sondern auch wesentliche qualitative Veränderungen.

# 3 Bisherige Ergebnisse

In diesem Abschnitt werden nun bisherige empirische Ergebnisse zu den Unterschieden in mathematischen Fähigkeiten, abhängig von den erhobenen individuellen Variablen, beschrieben. Die verwendeten Variablen werden im nächsten Abschnitt noch einmal genauer vorgestellt. Dabei werden ausschließlich Studien aus Deutschland verwendet, da diese Ergebnisse sich für verschiedene Bildungssysteme unterscheiden können. Eine Ausnahme stellen dabei die Untersuchungen zur verzögerten Studienaufnahme dar, da in diesem Fall keine Ergebnisse aus Deutschland vorliegen. Im letzten Teil werden dann Untersuchungen zum Erfolg von Studierenden in Klausuren im ersten Semester dargestellt.

## 3.1 Studiengang

Eine genauere Unterscheidung der verschiedenen Studiengänge innerhalb des MINT-Feldes wird im Allgemeinen nicht vorgenommen. Unterschieden werden in den meisten Fällen die sechs Interessenfelder nach Holland, sodass die MINT-Fächer zusammengefasst werden (Holland, 1997). Eine Ausnahme stellt die Untersuchung von Nagy (2007) dar. In dieser Untersuchung werden die Berufswünsche von $N = 3697$ Schülerinnen und Schüler aus Baden-Württemberg am Ende ihrer Schulzeit betrachtet, die im Rahmen des Projekts "Transformation des Sekundarschulsystems und akademische Karrieren" (TOSCA) befragt wurden. Dabei werden 33 Studiengänge unterschieden und die jeweiligen Gruppen neben ihren Interessenprofilen im Hinblick auf ihre kognitive Grundfähigkeit $F_k'$, basierend auf zwei Untertests aus dem "Kognitiven Fähigkeitstest für 4. bis 12. Klassen, Revision" (Heller & Perleth, 2000) sowie ihre Fachleistungen, kontrolliert bezüglich kognitiver Grundfähigkeit, in Mathematik, basierend auf Mathematikaufgaben zur voruniversitären Mathematik aus der Studie "Trends in International Mathematics and Science Study" (TIMSS) (Baumert, Bos & Lehmann, 2000), und Englisch untersucht. Insbesondere werden auch die drei Studiengänge Mathematik, Physik und Informatik unterschieden. Auf einer z-standardisierten Skala weisen dabei alle drei Studiengänge signifikant überdurchschnittliche kognitive Grundfähigkeiten auf, die zusammen mit Elektrotechnik die vier höchsten Werte darstellen. Der Studiengang Mathematik weist einen Mittelwert von $F_k' = 0,45$ aus, Physik von $F_k' = 0,77$ und Informatik besitzt $F_k' = 0,65$ als durchschnittlichem Wert. Die mathematischen Fachleistungen darüber hinaus zeigen ebenso mit $F_m' = 0,91$ für Mathematik und $F_m' = 0,71$ für Physik die beiden höchsten Werte und mit $F_m' = 0,39$ für Informatik einen signifikant überdurchschnittlichen Wert. Für die Fachleistung in Englisch ergeben sich hingegen keine signifikanten Unterschiede. Studierende im Lehramt werden für diese Untersuchung nicht gesondert angegeben.
Werden nun die Lehramtsstudierenden von den Nicht-Lehramtsstudierenden

unterschieden (Klusmann et al., 2009), so zeigt sich, dass die angestrebte Schulform der Lehramtsstudierenden einen wesentlichen Einfluss besitzt. Auf der Grundlage von $N = 1746$ Studierenden, wiederum im Rahmen der TOSCA-Untersuchung, konnten für die Abiturnote, kognitive Grundfähigkeit, Mathematikkompetenz sowie Englischkompetenz keine Unterschiede zwischen den Gymnasiallehramtsstudierenden und den Nicht-Lehramt-Studierenden an der Universität gefunden werden. Für die Unterschiede zu Grundschullehramtsstudierenden ergaben sich jedoch große Effekte bei die Abiturnote über mittlere Effekte in den Fachleistungen sowie kleine Effekte in der kognitiven Grundfertigkeit. Dabei wurden diese Effekte kontrolliert für Geschlecht und das Studieren mindestens einer Naturwissenschaft. Ebenso zeigten sich keine signifikanten Unterschiede in den gemessenen Persönlichkeitseigenschaften, hingegen gab es Unterschiede in den Dimensionen des beruflichen Interesses. Der Vergleich der Gesamtscores ergab für den Vergleich der Sportwissenschaftler mit den Sprachwissenschaftlern einen signifikanten Vorsprung der Sportwissenschaftler.
Ebenso besitzt im Vergleich innerhalb der Lehramtsstudierenden das jeweilige Fach einen wesentlichen Einfluss auf die kognitiven Fähigkeiten (Kaub et al., 2012). Dabei ergeben sich für $N = 227$ Lehramtsstudierenden signifikante Differenzen in Maßen der kognitiven Leistungsfähigkeiten. Als Maße der kognitiven Leistungsfähigkeit wurden dabei die Bereiche Allgemeinbildung, Denkfähigkeit, Worteinfall, räumliches Vorstellungsvermögen, Wahrnehmungsgeschwindigkeit sowie der durchschnittliche Prozentwert aus Horn (1983) verwendet. Die Studierendenden in Naturwissenschaften zeigten dabei insbesondere im Bereich des räumlichen Vorstellungsvermögens bessere Leistungen als die Studierenden in Geisteswissenschaften und Sprachwissenschaften. Studierende der Sportwissenschaften zeigten in der Denkfähigkeit und der Wahrnehmungsgeschwindigkeit Vorteile gegenüber den Studierenden der Geisteswissenschaften und zusätzliche im räumlichen Vorstellungsvermögen gegenüber den Studierenden der Geisteswissenschaften und der Sprachwissenschaften.
Eine weitere Untersuchung auf Grundlage von TOSCA-Daten untersucht gleichzeitig die Unterschiede zwischen Studierenden im Lehramt und Nicht-Lehramt sowie die Fächerwahl aus dem MINT-Bereich gegenüber Nicht-MINT-Fächern (Henoch et al., 2015). Dabei wurden nur angehende Gymnasiallehrkräfte untersucht und außerdem die Studiengänge der Nicht-Lehramt-Studierenden auf zum Lehramt korrespondierenden Fächer eingegrenzt. Als MINT-Studierende gelten alle Studierende, welche mindestens ein MINT-Fach studieren. Insgesamt ergibt sich dabei eine Stichprobengröße von $N = 1463$ Studierenden. Es zeigt sich, dass es zwischen den Studierenden im Lehramt und Nicht-Lehramt innerhalb der MINT-Fächer sowie außerhalb der MINT-Fächer jeweils keine Unterschiede in den kognitiven Fähigkeiten gab. Ein genereller Unterschied anhängig vom Lehramtsstudium ergibt sich bei dieser Betrachtung aus der geringeren Zahl von MINT-Studierenden inner-

halb der Lehramtsstudierenden.
Eine logistische Regression für die MINT-Studierenden mit der Variable "Lehramt" als abhängiger Variable zeigt, dass die gemessenen Persönlichkeitsmerkmale, Aufgeschlossenheit, Perfektionismus, Geselligkeit, Verträglichkeit sowie Neurotizismus, keinen signifikanten Einfluss auf die Variable besitzen, ebenso der sozioökonomische Hintergrund. Hingegen besitzen das Geschlecht sowie die Intressenseigenschaften signifikante Einflüsse. Ebenso besitzt die kognitive Fähigkeit mit einer Odds-Ratio von $\beta = 1,7$ einen signifikanten Einfluss auf die Entscheidung, Lehramt zu studieren, wobei die Studierenden mit höheren kognitiven Fähigkeiten eher einen Lehramtsstudiengang wählen. Dabei ist aufgrund einer nicht vorhandenen Korrelation zwischen dem Merkmal "Lehramt" und der kognitiven Fähigkeit jedoch eher von einem Supressoreffekt auszugehen als von einem direkten Einfluss.
Da in dieser Studie die MINT-Studiengänge gemeinsam untersucht wurden, ist jedoch unklar, ob sich die Befunde auch auf das einzelne Fach Mathematik übertragen lassen. Insbesondere, da die Unterscheidung zwischen MINT und Nicht-MINT bereits auf eine Abhängigkeit vom spezifischen Studienfach hinweist. So kann nicht ausgeschlossen werden, dass sich auch verschiedene MINT-Fächer verschieden verhalten. Insbesondere ist zu beachten, dass die verschiedenen Fächer verschiedene Zulassungsverfahren besitzen, sodass systematische Auswahleffekte entstehen.

## 3.2 Herkunftsbundesland

Durch das Institut zur Qualitätsentwicklung im Bildungswesen werden regelmäßig die Bildungsergebnisse in den Bundesländern untersucht. Dabei werden, entsprechend den PISA-Untersuchungen, alle drei Jahre die Kompetenzbereiche Mathematik sowie Naturwissenschaften und Deutsch, Englisch sowie Französisch im Wechsel untersucht. Mit dem Ländervergleich 2012 (Pant et al., 2013) wurde das erste Mal Mathematik untersucht, die nächste Erhebung findet entsprechend im Jahr 2018 statt.
Die Konzeption der Erhebung basiert auf den Bildungsstandards der Mittelstufe (Kultusministerkonferenz, 2004) und dem zugehörigen Kompetenzmodell. Dabei werden die fünf Leitideen "Zahl", "Messen", "Raum und Form", "Funktionaler Zusammenhang" und "Daten und Zufall" sowie die sechs allgemeinen mathematischen Kompetenzen "Mathematisch argumentieren", "Probleme mathematisch lösen", "Mathematisch modellieren", "Mathematische Darstellungen verwenden", "Mit symbolischen, formalen und technischen Elementen der Mathematik umgehen" und "Kommunizieren" für die Erhebung aus den Bildungsstandards übernommen, die drei Anforderungsniveaus hingegen werden durch ein sechsstufiges Kompetenzmodell ersetzt. Kompetenzen werden auf der PISA-Skala mit dem deutschlandweiten Mittelwert von 500 und einer Standardabweichung von 100 angegeben.

Insgesamt nahmen $N = 44584$ Schülerinnen und Schüler an der Untersuchung teil.
Auf der Gesamtskala liegen die Bundesländer Niedersachsen (495 Punkte), Hessen (495 Punkte) und Schleswig-Holstein (502 Punkte) auf einem Niveau mit dem Durchschnitt Deutschlands, während Nordrhein-Westfalen (486 Punkte) eine signifikant schlechtere Durchschnittskompetenz erzielt. In Bezug auf die durchschnittliche Kompetenz von Gymnasiasten, welche für Studienanfängerinnen und -anfänger als relevanter gelten kann, besitzen Nordrhein-Westfalen (581 Punkte) und Schleswig-Holstein (587 Punkte) mit dem deutschen Durchschnitt vergleichbare Leistungen und Niedersachsen (569 Punkte) sowie Hessen (570 Punkte) signifikant unterdurchschnittliche Leistungen. Der deutschlandweite Mittelwert der Gymnasiasten lag bei 586 Punkten.
Diese Ergebnisse zeigten sich im Wesentlichen unverändert bereits in der deutschen PISA-Erweiterungsstudie 2003 (Prenzel et al., 2005) zum Vergleich der Bundesländer: Für die Gesamtskala ergab sich ebenso nur für Nordrhein-Westfalen eine signifikante Abweichung vom Mittelwert, nach unten. Bei den Gymnasiasten zeigte Nordrhein-Westfalen 2003 entgegen 2012 noch eine durchschnittliche Kompetenz unter dem Bundesdurchschnitt, signifikante Abweichungen wurden nicht explizit angegeben. Die anderen drei Bundesländer zeigten an Gymnasien Leistungen, welche der Untersuchung von 2012 entsprechen.
Weitere Studien, welche die Leistungen von Schülerinnen und Schüler in den verschiedenen Bundesländern auf dem Niveau der Sekundarstufe I untersuchen, liegen nicht veröffentlicht vor. Auch Leistungsvergleiche in der Sekundarstufe II oder beim Verlassen der Schule wurden bisher nicht durchgeführt. Insbesondere gibt es auch keine Vergleiche der Bearbeitungen gemeinsamer Abituraufgaben, sowohl aus dem Pool der IQB als auch für die hilfsmittelfreien Aufgaben im Abitur, welche gemeinsam durchgeführt werden.
Neben den mathematischen Fähigkeiten in den verschiedenen Bundesländern, stellt auch das Studienortwahlverhalten der Studierenden einen wesentlichen Aspekt für die Leistungsmessung der Studienanfängerinnen und -anfänger dar. Hierbei zeigt sich, dass unter zwölf Gründen für die Wahl des Studienorts die Nähe zum Heimatort nur an der zehnten Stelle steht und von 41 % der Befragten als wichtig eingestuft wird (Schneider, Franke, Woisch & Spangenberg, 2017). Für Niedersachsen ergibt sich im Jahr 2015 dabei eine Quote von 57,1 % an Schulabgängerinnen und -abgängern, welche im eigenen Bundesland bleiben, eine im Vergleich eher hohe Quote, und liegt damit in einer Gruppe mit den weiteren großen Bundesländern in Deutschland (Sibbertsen & Stöver, 2017). Dabei zeigt sich gerade für Studierende mit ungünstigen Erfahrungen bezüglich des eigenen Studienerfolgs eine Tendenz ein Studium in räumlicher Nähe aufzunehmen (Helbig, Jähnen & Marczuk, 2017).

## 3.3 Geschlecht

An den Schulen erbringen Schülerinnen heute im Gegensatz zu Schülern bessere Noten und besitzen einen relativ höheren Anteil der Bildungsbeteiligung (Hannover & Kessels, 2011). Dabei stellt sich die Frage, ob diese Vorteile auch mit höheren Werten in schulunabhängigen Kompetenzmessungen einhergehen. Dazu sollen im Folgenden die Ergebnisse aus der PISA-Studie sowie der TIMSS-III-Studie zusammengefasst werden, welche für die letzten Jahre diese Frage untersucht haben.
Unterschiede in der Mittelstufe, abhängig vom Geschlecht, zeigt die PISA-Studie von 2012, in der Mathematik Hauptdomäne war (Prenzel et al., 2013). Dabei ergibt sich in Deutschland mit 14 Punkten Vorsprung für die Jungen gegenüber den Mädchen ein signifikanter Unterschied. Dieser Unterschied ist dabei größer als die Differenz im OECD-Durchschnitt, der 11 Punkte beträgt. Dennoch entspricht dies einem kleinen Effekt. Dabei überlagern sich die Verteilungen der beiden Gruppen wesentlich und die Unterschiede ergeben sich durch die besseren Ergebnisse der männlichen Spitzengruppe.
Für TIMSS-III (Baumert et al., 2000) ergibt sich, bei einem Mittelwert von 500 Punkten und einer Standardabweichung von 100 Punkten, in der mathematischen Grundbildung ein Vorsprung von 37 Punkten für die Männer, was einem mittleren Effekt entspricht. Diese Differenz vergrößert sich jedoch auf 62 Punkte für Schülerinnen und Schüler an allgemeinbildenden Gymnasien. Für die voruniversitäre Mathematik ergibt sich eine Differenz von 30 Punkten zwischen Schülerinnen und Schülern. Dabei ergibt sich diese Differenz nur in Leistungskursen, wo sie 25 Punkte beträgt, während sie auf Grundkursniveau nicht signifikant ist. Dabei ist zu beachten, dass 46 % der Schüler einen Leistungskurs in Mathematik besuchen, dem gegenüber aber nur 26 % der Schülerinnen.
Insgesamt ergeben sich kleine bis mittlere Vorsprünge in den Leistungen für Männer gegenüber Frauen. Dabei zeigen sich diese gerade in den Spitzengruppen. Diese entspricht damit eher den Studienanfängerinnen und -anfängern dieser Untersuchung.

## 3.4 Dauer des Schulbesuchs

Die Dauer der Schulzeit bis zum Abitur wurde ab dem Jahr 2007 in allen Bundesländern Deutschlands von dreizehn Jahren auf zwölf Jahre reduziert, nur die beiden Bundesländer Sachsen und Thüringen besaßen bereits vorher eine Schuldauer von zwölf Jahren. Mit dem ersten verkürzten Abiturjahrgang 2016 in Schleswig-Holstein und Rheinland-Pfalz hatten dann alle Bundesländer ein zwölfjähriges Abitur eingeführt. Inzwischen haben einige Bundesländer die Verkürzung auf eine Schulzeit von zwölf Jahren bis zum Abitur wieder zurückgenommen, zu diesem Bundesländern gehört insbesondere Niedersachen, in dem 2019 der letzte Abiturjahrgang die Schule nach zwölf Jahren

verlässt. Eine Übersicht und Einordnung der Reform in Deutschland gibt Homuth (2017). Im Weiteren sollen die Ergebnisse von Studien, welche sich mit Kompetenzmaßen auseinandersetzen, näher beschrieben werden.
Die Frage, welche Änderungen sich durch die Verkürzung der Schulzeit ergeben haben, wurde dabei nicht systematisch durch die Bundesländer untersucht. Die erreichten Abiturdurchschnittsnoten, aufgeteilt nach Schuldauer, wurde jedoch in vielen Fällen veröffentlicht. Ein Vergleich der durchschnittlich erzielten Abiturnoten in den Jahren mit doppeltem Abiturjahrgang zeigt jeweils nur geringe Unterschiede zwischen den beiden Gruppen: Im Abiturjahrgang Niedersachsen 2011 besaßen die Schülerinnen und Schüler mit dreizehnjährigem Abitur eine Durchschnittsnote von 2,55 und die Schülerinnen und Schüler mit einer Schulzeit von acht Jahren 2,58 als ihre Durchschnittsnote (Niedersächsisches Kultusministerium, 2011). Für den doppelten Abiturjahrgang 2013 in Nordrhein-Westfalen zeigt sich ein ebenso geringer Unterschied mit einem Abiturdurchschnitt von 2,41 für Schülerinnen und Schüler mit acht Jahren Schulzeit bzw. 2,44 für die Schülerinnen und Schüler mit neun Jahren Schulzeit (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2013). Für den ersten Doppeljahrgang in Hessen 2011 ergab sich ein Unterschied von 0,02 Notenpunkten für den Abiturdurchschnitt zu Gunsten der Schülerinnen und Schüler mit acht Jahren Gymnasialzeit. Dabei ergaben sich in keinem Fach Unterschiede von mehr als 0,1 auf der Notenskala (Hessisches Kultusministerium, 2011). Es zeigen sich also in keinem der Bundesländer wesentliche Unterschiede in der Abiturdurchschnittsnote[1][2].
Eine genauere Untersuchung der Ergebnisse in gemeinsamen Abschlussarbeiten in einem Doppeljahrgang liegt für Sachsen-Anhalt vor (Büttner & Thomsen, 2015). Hierfür wurden für ausgewählte Schulen die Leistungen in den zentralen Abschlussarbeiten in den Fächern Mathematik, Deutsch und Englisch von 724 Schülerinnen und Schülern untersucht: Es zeigt sich in Mathematik für die Schülerinnen und Schüler mit kürzerer Schulzeit ein signifikant schlechteres Abschneiden mit im Mittel 7,04 Punkten im Vergleich zu 7,75 Punkten. Dabei ergibt sich auch ein Interaktionseffekt mit dem Geschlecht. Für die Schüler beträgt der Unterschied ca. 1 Punkt, während für Schülerinnen eine Differenz von 0,5 Punkten vorliegt. Der Vergleich mit den weiteren untersuchten Fächern Englisch und Deutsch zeigt, dass sich dort ein wesentlich geringerer Effekt ergibt. In Deutsch verschwindet er vollständig, während in Englisch nur für die Schülerinnen eine schlechtere Leistung bei zwölfjährigem Abitur vorliegt.

[1]Hier und im Weiteren werden für Darstellungen bzgl. einzelner Bundesländer immer Niedersachsen, Nordrhein-Westfalen, Hessen sowie Schleswig-Holstein ausgewählt, da für diese Bundesländer die Studierendenanzahl eine statistische Auswertung zulässt (siehe dazu auch Abschnitt 6).

[2]Die Daten für einen Vergleich der Abiturienten im Doppeljahrgang liegen für Schleswig-Holstein nicht vor.

Im Unterschied zu den bisherigen auf Schulnoten basierenden Untersuchungen wurden in einer Untersuchung in Baden-Württemberg auch Leistungsunterschiede durch davon unabhängige Fähigkeitsmessungen durchgeführt (Hübner et al., 2017). Hierzu wurden die mathematischen Fähigkeiten, angelehnt an die PISA-Definition, in 21 Items erfragt. Teilgenommen haben 5210 Schülerinnen und Schüler aus dem Doppeljahrgang 2012 sowie den vorherigen und dem folgenden Jahrgang. Es zeigten sich keinerlei statistisch signifikanten Unterschiede in den Mathematikleistungen. Keine bzw. geringe Unterschiede zeigten sich ebenso für die Fächer Biologie und Physik, ein statistischer Vorsprung für die Schülerinnen und Schüler mit dreizehnjährigem Abitur ergab sich hingegen in Englisch.
Auch in Hamburg wurden im Rahmen der Untersuchung "Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der gymnasialen Oberstufe" (KESS) für den Abiturjahrgang 2012/13 Vergleiche mit dem Schuljahrgang 2005/06 der Untersuchung "Aspekte der Lernausgangslage und der Lernentwicklung" (LAU) gezogen (Vieluf, Ivanov & Nikolova, 2011). In beiden Untersuchungen wurden die mathematischen Fähigkeiten der Schülerinnen und Schüler unter Anderem vor Beginn des siebten Schuljahrgangs sowie am Ende der Oberstufe mit den nahezu identischen Instrumenten erfasst. Dabei wurden im Rahmen der Konzeption der TIMSS-Untersuchung sowohl die "Mathematische Grundbildung" als Vollerhebung und "Voruniversitäre Mathematik" an einer Stichprobe erhoben. Am Ende der Oberstufe zeigt sich für die mathematische Grundbildung mit einer Effektstärke von $d=0,14$ ein kleiner Effekt für die Schülerinnen und Schüler mit neunjähriger Schulzeit, für die voruniversitäre Mathematik zeigt sich mit $d=0,05$ ein kleiner Effekt in die andere Richtung. Für die ebenso untersuchten Englischleistungen zeigen sich hingegen Vorsprünge der Schülerinnen und Schüler mit kürzerer Schulzeit, während in Naturwissenschaften kein signifikanter Unterschied festgestellt wurde. Weitere Analysen zeigen dabei, dass die Effekte auch bei Kontrolle weiterer personenbezogener Hintergrundvariablen erhalten bleiben.
Die bisherigen vorgestellten Untersuchungen vergleichen jeweils Gruppen von Schülerinnen und Schülern am Ende ihrer Schulzeit unabhängig von folgenden Entscheidungen für ein Studium bzw. Studiengang. Eine Untersuchung zu Unterschieden von Studienanfängerinnen und -anfängern liegt für die Duale Hochschule Baden-Württemberg Mannheim vor (Derr & Hübl, 2014; Derr, Hübl & Ahmed, 2013). Hier wird ein Online-Test zur Diagnose der Eingangsvoraussetzungen verwendet, der neunzig Fragen zu Mittelstufen- und Oberstufenmathematik enthält. Auf Grundlage der drei Jahrgänge 2011 bis 2013, mit einer Beteiligung von jeweils ca. 80% der Anfängerinnen und Anfänger, wurden die Leistungen in Abhängigkeit von der Schuldauer für die vier Bundesländer Baden-Württemberg, Rheinland-Pfalz, Hessen und Nordrhein-Westfalen verglichen. Dabei konnte für keines der Bundesländer ein Effekt gefunden werden, auch nicht über alle Bundesländer hinweg.

Ein ähnliches Ergebnis zeigt eine Untersuchung zu Eingangsvoraussetzungen der Studierenden an einer Universität in Nordrhein-Westfalen im Doppeljahrgang 2013 (Kühn, 2014). Hier wurden unabhängig vom gewählten Studiengang alle Studienanfängerinnen und -anfänger befragt, sodass sich eine Teilnehmendenanzahl von $N = 1460$ ergab. Neben einigen weiteren Konstrukten wurde auch ein Item zur Selbsteinschätzung in Mathematik verwendet. Es zeigten sich dabei keine signifikanten Unterschiede für Schülerinnen und Schüler mit unterschiedlichen Schuldauern.

Unter Verwendung der PISA-Daten für Deutschland aus den Jahren 2000 bis 2015 zeigte sich in Mathematik ein kleiner Effekt der Kompetenz zu Gunsten der Schülerinnen und Schülern mit zwölf Jahren Schulzeit von $d = 0,075$, die beiden anderen Kompetenzbereiche zeigten die gleiche Effektstärke (Andrietti, 2016). Dabei wurde die Zugehörigkeit für jedes Bundesland durch Erhebungszeitraum und Zeitpunkt der Umstellung ermittelt. Tiefere Auswertungen zeigen, dass die Unterschiede in Mathematik sich unabhängig vom Geschlecht ergeben und Schülerinnen und Schüler, deren Eltern einen hohen Bildungsgrad besitzen, von der Verkürzung profitieren (Andrietti & Su, 2018).

Neben den beschriebenen Untersuchungen bezüglich veränderten Kompetenzen von Schulabgängerinnen und -abgängern sind auch Auswirkungen auf die Auswahl des folgenden Studienfachs möglich. Dies wurde für den Doppeljahrgang in Sachsen-Anhalt durch eine Briefumfrage einundzwanzig Monate nach dem Abitur mit einer Teilnehmerzahl von $N = 724$ Abiturientinnen und Abiturienten durchgeführt (Meyer & Thomsen, 2016). Dabei zeigte sich kein Effekt auf die Aufnahme eines Studiums im MINT-Bereich. Eine vertiefte Analyse auf Fächerebene hingegen ergab, dass sich für männliche Studenten die Wahrscheinlichkeit für einen Studienbeginn in Mathematik um 15 % reduzierte.

Zusammenfassend zeigen die Untersuchungen kein einheitliches Bild in den Kompetenzunterschieden für Schülerinnen und Schüler mit unterschiedlichen Schuldauern. Für die nicht auf Schulnoten bezogenen Kompetenzmessungen ergaben sich eher Vorteile für die Schülerinnen und Schüler mit verkürzter Schulzeit. In jedem Fall sind nur Effekte mit kleiner Ausprägung vorhanden und zum Teil nur in Kombination mit weiteren Merkmalen zu finden. Dabei ist insbesondere zu beachten, dass die Vergleiche mit Stichproben verschiedener Jahrgänge, wie es die Untersuchung aus Hamburg und die Untersuchung zu den PISA-Daten sind, die größten Effektstärken aufweisen. In diesen Vergleichen werden also auch Auswirkungen von weiteren Veränderungen in Schule und Unterricht beinhaltet sein, welche einen Beitrag zu den Effekten leisten könnten. Insgesamt können also nur kleine Effekte in den mathematischen Fähigkeiten im Vergleich der verschiedenen Schulzeiten für die Studienanfängerinnen und -anfänger erwartet werden.

### 3.5 Verzögerter Studienbeginn

Untersuchungen zur Verzögerung des Studienbeginns gibt es im deutschen Bereich wenig. Insbesondere für die Frage der Fachkompetenzen von Studienanfängerinnen und -anfängern mit einem verzögerten Studienbeginn fehlen Untersuchungen. Daher werden in diesem Abschnitt auch Untersuchungen aus den USA zu diesem Themenbereich dargestellt. Auch in den USA werden im Wesentlichen die Gründe für die Verzögerung betrachtet, die auch in Deutschland erhoben wurde. Dabei werden in Fall der Untersuchungen in den USA auch fachliche Unterschiede in der Schule als Begründungen für die verzögerte Studienaufnahme verwendet. Dabei liegt der Fokus der Forschung eher auf den Ursachen der Verzögerung, insbesondere im Bezug auf soziale Ungleichheiten, als auf den Folgen.

In der Untersuchung des Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (Schneider et al., 2017) wurden 2015 Hochschulzugangsberechtigte zu zwei Zeitpunkten befragt. Insgesamt ergab sich eine Stichprobe von $N = 8953$ Teilnehmenden, welche sowohl am ersten Zeitpunkt, ca. ein halbes Jahr vor dem Schulabschluss sowie im Dezember des gleichen Jahres befragt wurden. Dabei wurden in dieser Untersuchung nicht nur Studierende befragt, sondern alle Hochschulzugangsberechtigten. Dadurch ergibt sich eine Quote von 70 % der Hochschulzugangsberechtigten, welche direkt ein Studium oder eine andere Tätigkeit aufgenommen haben. Werden nur die Studierenden betrachtet, ergibt sich eine wesentlich geringere Quote, die bei 35 % liegt (Middendorff et al., 2013).

Unter dreizehn erhobenen Gründen für die Verzögerung gaben die Studienberechtigten unter möglicher Mehrfachnennung als den häufigsten Grund mit 45 % "nach Schule erst einmal eine Pause" an sowie als dritthäufigsten Grund mit 28 % "wollte zunächst längere Zeit ins Ausland". Als die beiden wichtigsten berufsbezogenen Begründungen wurden "unschlüssig über Werdegang" sowie "Zulassungsbeschränkung für das gewählt Fach" mit 38% bzw. 20% genannt. Weiterhin wurden zu 21% mit "vorher Geld verdienen" finanzielle Gründe genannt.

Wie beschrieben sollen trotz der Unterschiede zwischen Deutschland und den USA nun einige Ergebnisse zum verzögerten Studienbeginn aus den USA beschrieben werden. Dabei ist zu bedenken, dass für das Jahr 2003/04 eine Verzögerung für 17 % der Studienanfängerinnen und -anfängern in den USA angegeben wird (Hoe, 2014) und damit das Phänomen wesentlich geringer als in Deutschland (Middendorff et al., 2013) ausfällt. Dies macht deutlich, dass sich die Ergebnisse nicht direkt auf Deutschland übertragen lassen.

Für $N = 129990$ Studierende einer nationalen Studie der USA wurden soziale sowie schulbezogene Leistungsdaten erhoben und auf ihren Einfluss für einen möglichen verzögerten Studienbeginn untersucht. Dabei zeigten sich als signifikante Einflussfaktoren auf die verzögerte Studienaufnahme die Kurswahl höherer mathematischer Veranstaltungen in der high school sowie hohe

Ergebnisse in standardisierten Leistungstest zur Studieneignung. In dieser Analyse zeigte sich ebenso, dass Frauen eher den Studienbeginn verzögern sowie finanziell unabhängige Studierende seltener den Studienbeginn verzögern. Daneben ergab die deskriptive Analyse, dass Studierende mit verzögertem Beginn schlechtere Schulnoten besitzen sowie eher aus Familien mit Einkommen im unteren Quartil stammen. Die anderen drei Quartile zeigten hingegen keine Unterschiede.
Zu sehr ähnlichen Ergebnissen kommen auch weitere Studien über verschiedene Zeitpunkte (Niu & Tienda, 2013; Roksa & Velez, 2012; Bozick & DeLuca, 2005). In diesen Studien zeigt sich jeweils, dass das Einkommen der Eltern und die schulischen Leistungen bzw. Ergebnisse standardisierter Test wesentliche Erklärungsfaktoren für die Verzögerung des Studienbeginns darstellen. Dabei zeigt die älteste der drei Studien mit N=11336 Teilnehmenden, dass Studierende im höchsten Punktequartil nur halb so oft ihren Studienbeginn verzögern wie Studierende im unteren Quartil. Im Vergleich des untersten und des obersten Quartils für den sozioökonomischen Hintergrund ergibt sich sogar ein Abfall auf 30% für das Verzögerungsrisiko. Dabei sind hier die Unterschiede für alle Quartile signifikant, während für die Leistungsdaten nur der Unterschied der beiden extremen Quartile signifikante Unterschiede erzielt.
Insgesamt zeigt sich also, dass der verzögerte Studienbeginn in allen Fällen einen deutlichen Zusammenhang mit schulischen Ergebnissen bzw. Testergebnissen besitzt, auch wenn sich im Einzelnen Unterschiede ergeben. Die ebenso gefundenen sozialen Zusammenhänge werden hier nicht weiter betrachtet.

## 3.6 Kursform

Unterschiede in den Kompetenzen von Schülerinnen und Schülern am Ende ihre Schullaufbahn in Abhängigkeit der Kurswahl in der Oberstufe lassen sich auf verschiedene Faktoren zurückführen. Dabei zeigen sich bereits vor der Aufteilung in die verschiedenen Kursformen, wie Grundkurs oder Leistungskurs bzw. Ähnliches, Differenzen zwischen Schülerinnen und Schülern, sowohl in den mathematischen Fähigkeiten als auch im Interesse, die zu unterschiedlichen Kursformwahlen führen. Darüber hinaus unterscheiden sich die Kompetenzverläufe in den Kursformen, sodass sich vergrößerte Unterschiede ergeben.
Im Rahmen der TIMSS-III-Untersuchung (Baumert et al., 2000) wurden die Kurswahlmotive von Schülerinnen und Schülern für die Entscheidung eines Mathematikleistungskurs untersucht. Dabei ergaben sich mit Hilfe einer Hauptkomponentenanalyse zwei Faktoren: Der erste Faktor erklärt ungefähr 30 % der Gesamtvarianz und beschreibt Kompetenzentfaltung, Interesse sowie die Berufsperspektive. Der zweite Faktor beschreibt soziale Faktoren wie den Kontakt zu Lehrkräften oder Mitschülerinnen und -schülern. Dieser Fak-

tor beschreibt knapp 20 % der Varianz.
In Bezug auf die Mathematikleistungen zeigen sich deutliche Unterschiede zwischen den Kursniveaus am Ende der Schulzeit: Dabei ergibt sich für die Schülerinnen und Schüler im Leistungskurs mit d=0,9 ein starker Effekt. Interessant ist dabei weiterhin, dass die Schülerinnen und Schüler, welche als weiteren Leistungskurs Physik gewählt haben, einen signifikanten Vorsprung im Vergleich zu anderen Leistungskursen besitzen. Ebenso zeigt sich, dass die Kurswahlen mit einer Verringerung der Varianz in den Leistungskursen einhergehen. Hierbei ist zu bemerken, dass seit der Durchführung der TIMSS-III-Untersuchung die Schulsysteme in Deutschland wesentlich geändert haben, sodass auf der einen Seite die Kurswahlen verändert wurden und auf der anderen Seite auch die Unterscheidung der Stundenanzahl heute eine andere ist.
In der TOSCA-Untersuchung wurden ebenfalls die Schülerinnen und Schüler in ihrem letzten Schuljahr auf ihre Mathematikleistungen in voruniversitärer Mathematik im Rahmen der Konzeption von TIMSS untersucht, unterschieden in Abhängigkeit des Gymnasialtyps und nach der Kursform. Dabei ergibt sich am allgemeinbildenden Gymnasium wiederum ein Unterschied zwischen den beiden Kurstypen, der über einer Standardabweichung liegt, also wiederum ein starker Effekt. Auch an allen anderen Gymnasialformen ergeben sich deutlich Mittelwertunterschiede zwischen Leistungskursen und Grundkursen, wobei sowohl der Haupteffekt der Gymnasialform als auch der Interaktionseffekt signifikant werden. In einer linearen Regressionsanalyse unter Verwendung der Schulform, des Geschlechts, des sozioökonomischen Status, einer möglichen Klassenwiederholung, sowie Vorwissenskomponenten, zeigt der Kurstyp einen Vorsprung von 60 Punkten auf der TIMSS-Skala für die Schülerinnen und Schüler im Leistungskurs.
Entsprechende Ergebnisse zeigt auch die LISA-6-Studie (Leucht, Kampa & Köller, 2016), in der Schülerinnen und Schüler in ihrem letzten Schuljahr aus Schleswig-Holstein untersucht werden, Grundlage ist dabei der Test des Nationalen Bildungspanels. Da hier alle Schülerinnen und Schüler auf den allgemeinbildenden Gymnasien auf vierstündigem erhöhtem Anforderungsniveau unterrichtet wurden, ergeben sich nur für berufliche Gymnasien verschiedene Kurswahlen: dreistündig auf grundlegendem Anforderungsniveau und fünfstündig auf erhöhtem Anforderungsniveau. Dabei zeigt sich zwischen den allgemeinbildenden Gymnasien und den eA-Kursen der berufsbildenden Gymnasien ein vernachlässigbarer Effekt, beide Gruppen besitzen allerdings mit einer Effektstärke größer 1 einen starken Vorsprung gegenüber den gA-Kursen der beruflichen Gymnasien. Bei der Interpretation ist zu bedenken, dass die berufsbildenden Gymnasien nicht weiter nach ihrer Art aufgespalten werden und sich die Anteile der Leistungskurswahlen in Mathematik zwischen den Formen wesentlich unterscheiden.
Für grundständige Gymnasien in Hamburg zeigen die Untersuchungen KESS und LAU, dass sich, unabhängig von zwölfjähriger oder dreizehnjähriger

Schulzeit, im Vergleich der jeweiligen Kursformen am Ende der Schulzeit große Unterschiede ergeben. Diese ergeben sich dabei sowohl in den Bereichen der mathematischen Grundbildung als auch der voruniversitären Mathematik im Rahmen der TIMSS-Konzeption. Einzig der Unterschied zwischen eA-Kurs und gA-Kurs im Vergleich zwischen Leistungskurs und Grundkurs im Bezug auf die voruniversitäre Mathematik fällt auf hohem Niveau etwas geringer aus.
Die Eingangsvoraussetzungen von Schülerinnen und Schülern in Abhängigkeit der Kurswahl wurde an einem niedersächsischen Fachgymnasium mit dem Zweig Wirtschaft untersucht (Warwas, 2008). Dabei zeigte sich erwartungsgemäß, dass für N=135 Schülerinnen und Schüler die Gruppe, die einen Kurs mit erhöhtem Anforderungsniveau anstrebt, höhere Leistungen in mathematischer Grundbildung erzielen. Die geplante Kursform erklärt dabei ungefähr 10 % der Varianz zwischen den Schülerinnen und Schülern.
Insgesamt lässt sich also feststellen, dass die Wahl der Kursform im Mathematikunterricht über alle Untersuchungen hinweg mit verschiedenen Schulsystemen und Schulformen einen starken Effekt auf die mathematischen Fähigkeiten besitzt. Dieser ist dabei auch bereits vor der Oberstufe zu erkennen, also vor der Entscheidung für die Kursform.

## 3.7 Schulzensuren

Im Folgenden werden einige Ergebnisse zu den Zusammenhängen zwischen Schulzensuren und Leistungen in standardisierten Tests beschrieben werden, in diesem Fall werden in jeder Untersuchung die TIMSS-Fragebögen verwendet. Dabei liegt der Fokus auf den globalen Zusammenhängen und nicht auf individuellen Faktoren, die zu Verzerrungen in der Beurteilung einzelner Schülerinnen und Schüler führen. Dazu gibt (Holmeier, 2012) eine kurze Übersicht. Die Untersuchungen sind, mit Ausnahme der TIMSS-III-Studie, dabei im Rahmen von Untersuchungen zu Wirkungen des Zentralabiturs entstanden, von dem sich eine Erhöhung der Einheitlichkeit von Bewertungsmaßstäben versprochen wird.
Als Erstes wurde in der TIMMS-III-Studie (Baumert et al., 2000) der Zusammenhang zwischen den Testergebnisse und Schulnoten der Schülerinnen und Schüler untersucht. Dabei ergaben sich global Korrelationen um $r = 0,4$ zwischen Zensuren und Testleistungen. Es zeigten sich dabei klare Referenzgruppeneffekte: Schülerinnen und Schüler an im Mittel schwächeren Schulen erreichten bei gleichen Schulnoten wesentlich schwäche Durchschnittswerte im Mathematiktest.
Ein Vergleich der beiden Bundesländer Hessen und Bremen (Holmeier, 2012) ergibt für die Jahre 2007 bis 2009 Unterschiede in den Korrelationen zwischen Mathematikleistung und Abiturnote. Die Korrelationen liegen ja nach Jahr und Bundesland für Leistungskurse zwischen 0,4 und 0,6. Die Schwankungen in den Grundkursen sind wesentlich größer, sodass sich Korrelationen

zwischen 0,2 und 0,55 ergeben. Dabei lassen sich keine klaren Muster für die Stärke des Zusammenhangs finden.
Für einen Vergleich der zentralen Abiturklausuren in Nordrhein-Westfalen (Kahnert, 2014) mit dem TIMMS-Leistungstest ergibt sich eine Rangkorrelation von $\rho = 0,5$, dessen Höhe also dem obigen Ergebnis entspricht. Dabei zeigt sich weiterhin, dass eine gemeine Modellierung der TIMMS-Aufgaben und der Abituraufgaben mit dem Rasch-Modell möglich ist, also von Messungen der gleichen Kompetenz ausgegangen werden kann.
Im Vergleich der beiden Bundesländer Hamburg und Baden-Württemberg (Neumann et al., 2009), auf Grundlage von Daten aus den Erhebungen LAU und TOSCA, zeigt sich, dass in Mathematik die Durchschnittsfachnoten im ersten Halbjahr des dreizehnten Schuljahrs keine Differenz zwischen den Bundesländern aufweisen und es im Abiturdurchschnitt eine Differenz mit kleinem Effekt gibt, dabei besitzen die Schülerinnen und Schüler aus Baden-Württemberg etwas bessere Durchschnittsnoten. Gleichzeitig zeigen diese Schülerinnen und Schüler allerdings im Leistungskurs eine mit einer Effektstärke von $d = 0,76$ und im Grundkurs mit $d = 1,01$ bessere durchschnittliche Leistung im TIMMS-Test. Weiterhin wurden mit Hilfe von linearen Regressionen sowohl die Schulnote im ersten Halbjahr der Klasse 13 als auch in der zentralen Abschlussprüfung untersucht. Dabei zeigt sich, dass die Differenz, die sich aufgrund des Bundeslandes ergibt, für die Schulnoten im Unterricht wesentlich größer ist als es für die Abiturprüfungen der Fall ist. Für die Abiturprüfungen verschwand dieser Effekt sogar unter Kontrolle von Referenzgruppeneffekten. In beiden Analysen zeigten sich Varianzaufklärungen innerhalb der Schulen von ca. $R^2 = 0,4$.
Insgesamt ergeben sich also zwischen den Schulzensuren und den Ergebnissen der Leistungsmessungen Korrelationen, die bei r=0,4 bis r=0,6 liegen und damit von mittlerer bis starker Größe sind.

## 3.8 Prädiktion von Studienerfolg

Als Letztes sollen verschiedene Befunde zur Vorhersage von Studienerfolg im ersten Semester berichtet werden. Dabei spielt insbesondere die Abiturnote eine entscheidende Rolle. Es werden zunächst Metaanalysen verschiedener Studiengänge vorgestellt, welche die Prädiktion aufgrund verschiedener Maße beschrieben, anschließend werden diese anhand einzelner Untersuchungen aus der Mathematik und Physik bzw. Chemie genauer betrachtet. Als Letztes werden Ergebnisse aus Studien zu den Prädiktoren, die hier erhoben wurden, dargestellt.
In einer Metaanalyse zur Prädiktionskraft von Schulabschlussnoten aus fünf europäischen Ländern wird der Zusammenhang mit Studiennoten als Erfolgskriterium untersucht (Trapmann, Hell, Weigand & Schuler, 2007). Es werden insgesamt 26 Einzelstudien für die Untersuchung verwendet, wobei diese aus verschiedenen Studiengängen stammen. Ein erheblicher Anteil der Studien-

gänge sind medizinische Studiengänge. Dabei ergeben sich eine prädiktive Validität von $\rho = 0,52$ für die Schulabschlussnote. Für die Verwendung von Einzelnoten ergeben sich wesentlich geringere Werte, die bei Werten bis $\rho = 0,4$ liegen. Dies zeigt sich auch im Vergleich zu den sogenannten "big five" der Persönlichkeitsmerkmale (Trapmann, Hell, Hirn & Schuler, 2007), welche in einer ähnlichen Studie mit 52 Studien nur einen Wert von $\rho = 0,27$ aufweisen.
Damit zeigen sich die Schulabschlussnoten auch im Vergleich mit spezifischen Eignungstests verschiedener Studiengänge als ähnlich bis höher in ihrer Prädiktion von Studiennoten (Hell, Trapmann & Schuler, 2007). Für 36 verwendete Einzelstichproben ergab sich insgesamt mit $\rho = 0,48$ eine prädiktive Validität in Höhe des Werts für die Schulabschlussnoten. Auch hier beziehen sich die Analysen zu einem großen Teil auf medizinische Studiengänge.
Im Vergleich zeigt sich weiterhin für die Studiengänge Medizin, Jura, Ingenieurswissenschaften und Wirtschaftswissenschaften mit insgesamt $N = 395$ Teilnehmenden (Gold & Souvignier, 2005), dass fachspezifische Wissenstests und Fachinteresse keine zusätzlichen Beiträge für die Erklärung der Examensnote liefern.
Insgesamt ergibt sich also, dass die Abiturnote als Einzelprädiktor die höchste Prädiktionskraft für den Erfolg im Studium besitzt. Insbesondere zeigen spezifische Eingangstests keine wesentlich besseren Erklärungsstärken. Die letzte Studie weist ebenso darauf hin, dass die nicht-kognitiven Faktoren einen eher geringeren Einfluss besitzen.
Es stellt sich die Frage, ob sich die beschriebenen Ergebnisse auch auf die hier betrachteten Studiengänge übertragen lassen. Dazu werden nun Untersuchungen zu Mathematik, Physik und als ein weiteres naturwissenschaftliches Studienfach Chemie dargestellt. In Mathematik zeigten sich, dass für $N = 182$ Studierende die mathematische Kompetenz bezogen auf "Analysis 1" vor der Vorlesung sowie die Abiturdurchschnittsnote einen signifikanten Einfluss auf das Bestehen in der Veranstaltung "Analysis 1" besitzen (Rach & Heinze, 2013, 2017). Dabei erklärt die mathematische Kompetenz knapp 30 % der Varianz und die Abiturnote zusätzliche 9 %. Die außerdem erhobenen Variablen Interesse an Mathematik, das mathematikbezogene Selbstkonzept sowie die extrinsische Studienmotivation lieferten keine zusätzlichen Beiträge.
Für Physik (Sorge, Petersen & Neumann, 2016) zeigt sich bei $N = 77$ Studierenden im Vergleich der drei Leistungsmaße Abiturnote, letzte Physiknote und einem Vorwissenstest in Physik, dass die Verwendung der beiden Schulvariablen signifikante Beiträge innerhalb einer logistischen Regression zum Klausurbestehen liefern. Dieses Modell konnte 87 % der Fälle korrekt zuordnen.
Für $N = 165$ Studierende in Chemie (Freyer, Epple, Brand, Schiebener & Sumfleth, 2014) zeigt eine Regressionsanalyse für die Punktzahl in der Klausur, dass die Variablen Abiturnote, Vorwissen, Fähigkeit im schlussfolgern-

den Denken, Fachinteresse sowie der Studiengang (Lehramt) einen signifikanten Beitrag liefern, während die Variable Wunschfach als einzig erhobene Variable keinen signifikanten Beitrag liefert. Dabei werden insgesamt jedoch nur 28,5 % der Varianz aufgeklärt, wobei das chemische Fachwissen mit 13,3 % den höchsten Einzelbeitrag liefert und die Abiturnote nur gut 7 % zusätzliche Varianzaufklärung beiträgt.
Für Mathematik und auch Physik zeigen also auch hier die Studien, dass die Abiturnote den größten Einfluss auf den Studienerfolg, konzeptualisiert als Klausurnote oder Bestehen der Klausur, liefern. Die nicht leistungsbezogenen Variablen liefern in Mathematik keine zusätzlichen Beiträge und in Chemie auch nur einen kleinen Beitrag. Dies bestätigt insgesamt die oben an größeren Gruppen beschriebenen Befunde für Mathematik bzw. naturwissenschaftliche Studiengänge.
Im Folgenden werden nun noch einige Befunde zu einzelnen der erhobenen Variablen berichtet. Im Vergleich von insgesamt $N = 4860$ Studierenden der Universität Konstanz wurden Kohorten verglichen, welche ein zwölfjähriges bzw. dreizehnjähriges Abitur gemacht haben (Dörsam & Lauber, 2015). Dabei zeigte sich bei zwei verglichenen Kohorten, dass in der ersten keine Unterschiede auffindbar waren, während für den zweiten Vergleich eine Differenz in Klausurergebnisse zu Gunsten der verkürzten Schuldauer mit einer Effektstärke von $d = 0,2$ vorlag. Dies beschreibt also insgesamt eine kleine Differenz. In Bezug auf das Geschlecht zeigt sich in den Abbruchquoten der Jahre von 1992 bis 2007 für Frauen Abbruchquoten, die gute 10 % über den Abbruchquoten für Männer liegen (Dieter, 2012), wobei hier keine Kontrolle bezüglich weiterer Variablen durchgeführt wurde. Auch im Bezug auf die Leistungskurswahlen (Fries, 2002) in der Oberstufe ergibt sich, dass die Studierenden, welche ein zu ihrem Fach passenden Leistungskurs gewählt haben, bessere Leistungen an der Universität erzielen als die Studierenden, deren Leistungskurswahlen nicht zu ihrem Studiengang passen.
Im Bezug auf die verzögerte Studienaufnahme ergibt eine eigene Untersuchung (Halverscheid & Pustelnik, 2013) für den Jahrgang 2011, deren Stichprobe nicht Teil dieser Untersuchung ist, dass für die Vorlesung "Differential- und Integralrechnung 1" die Studierenden mit verzögerter Studienaufnahme mit einer kleinen Effektstärke von $d = 0,21$ bessere Klausurergebnisse erzielen bei gleichen Eingangstestergebnissen. Ebenso zeigen Studierende mit dreizehnjähriger Schulzeit hier mit $d = 0,15$ schwächer Klausurergebnisse. Dabei wird der, bis auf kleine Veränderungen, gleiche Test wie in dieser Untersuchung verwendet.
Insgesamt zeigen die Studien, dass der stärkste Prädiktor für den Erfolg im ersten Studiensemester die Abiturnote ist. Die weiteren Beiträge fachspezifischer Leistungstests bzw. Eignungstest schwanken erheblich bis zu einer Höhe, die der Abiturnote entsprechen kann. Weitere Merkmale besitzen nur geringe zusätzliche Erklärungskraft. Für die zusätzlichen Prädiktoren, zeigen sich geringe Effekte auf den Studienerfolg.

# 4 Forschungsfragen

Aus den oben beschriebenen theoretischen Forschungsrahmen sowie den bisherigen Befunden ergeben sich nun für zwei Bereiche getrennt die Forschungsfragen. Der erste Teil beschreibt, welche fachmathematischen Voraussetzungen für ein erfolgreiches Studium die Studienanfängerinnen und -anfänger zu Beginn ihres Studiums mitbringen. Dabei wird untersucht, von welchen der untersuchten Variablen diese abhängen. Im zweiten Abschnitt soll dann der Erfolg, im Sinne des Bestehens von Klausuren im ersten Semester, in Zusammenhang gebracht werden mit den vorher erhobenen Eingangsvoraussetzungen. In beiden Bereichen wird zusätzlich der Frage nachgegangen, ob eine Betrachtung spezifischer mathematischer Gebiete weiteres Erklärungspotential bietet.

## 4.1 Testergebnisse

Da in dieser Arbeit der Erfolg im Sinne des Bestehens von Klausuren untersucht wird, haben die beschriebenen bisherigen Untersuchungen aufgezeigt, dass der erklärungsstärkste Prädiktor die mathematische Leistung vor dem Studium darstellt bzw. die Abiturnote. Daher stellt sich die Frage, wovon diese mathematischen Fähigkeiten abhängen. Dazu werden im Einzelnen verschiedene Gruppen von Studierenden untersucht, welche aufgrund individueller Merkmale vergleichen werden. Hierfür werden die untersuchten Fragen offen formuliert, da trotz der vorgestellten Ergebnisse zwar Hypothesen über die Wirkungen aufgestellt werden können, gleichzeitig aber sich in vielen Fällen die Stichprobenzusammensetzung wesentlich unterscheidet, das Alter verschieden ist oder auch verschiedene Ergebnisse gefunden wurden. Dennoch sollen kurze Einschätzungen als Hypothesen, abgeleitet von den bisherigen Ergebnissen, angegeben werden. Anschließend wird darüberhinaus der Frage nachgegangen, ob es für verschiedene mathematische Bereiche verschiedene Einflüsse gibt, insbesondere, ob diese abhängig sind von der zeitlichen Stellung im Schulcurriculum.
Als Erstes stellt sich die Frage, ob es zwischen den verschiedenen Erhebungszeitpunkten Unterschiede zwischen den Studienanfängerinnen und -anfängern in ihren Testergebnissen gibt. Es ist zu erwarten, dass sich keine Änderungen zeigen werden, sodass insbesondere eine Untersuchung der vier Zeitpunkte als eine Grundgesamtheit gerechtfertigt ist.

(F1) Gibt es zu den verschiedenen Erhebungszeitpunkten Unterschiede in den durchschnittlichen Testergebnissen?

Im Bezug auf die verschiedenen Studiengänge ergibt sich aufgrund der vorgestellten Studie von Nagy (2007) die Erwartung, dass die Studierenden in Informatik bezogen auf ihre Mathematikfertigkeiten etwas schwächere Leistungen erbringen werden als die Studierenden in Mathematik und in Physik. Dies spiegelt auch der mathematische Anspruch der Vorlesungen des ersten Semesters für die Studiengänge wider. Im Bezug auf die Studierenden im Lehramt lässt sich kein klares Bild ableiten. Hierbei stellt sich außerdem die Frage, wie stark sich der Selektionseffekt für die Auswahl eines der vier Studiengänge auf die mathematischen Fähigkeiten darstellt, also wie der Vergleich zwischen den Studienanfängerinnen und -anfängern mit den Schülerinnen und Schülern ausfällt.

(F2) Gibt es verschieden Mittelwerte der Testergebnisse in Abhängigkeit vom gewählten Studiengang?

(F2b) Wie schneiden die Schülerinnen und Schüler im Kurs mit erhöhtem Anforderungsniveau am Ende ihrer Schulzeit im Vergleich mit den Studierenden, unabhängig vom gewählten Fach ab?

(F2c) Wo sind die Schülerinnen und Schüler im Verhältnis zu den verschiedenen Studiengängen einzuordnen?

Im Vergleich der Bundesländer zeigten sich für die Fünfzehnjährige Unterschiede in den Leistungen zwischen den Schülerinnen und Schülern, auch im Vergleich auf die jeweiligen Spitzengruppen. Hierbei stellt sich die Frage, wie stark der regionale Einfluss der Universität im Verhältnis zu diesen Ergebnisse ist, die auch vier Jahre vor dem Schulabschluss liegen.

(F3) Welche Unterschiede im durchschnittlichen Testergebnis besitzen die Studierenden in Abhängigkeit ihres Herkunftsbundeslands?

Für den Vergleich von Männern und Frauen zeigten sich, wie im Abschnitt 3.3 vorgestellt, konstante Vorteile für die Männer. Dabei lag jeweils kein großer Abstand zwischen den beiden Geschlechtern vor. Da dieser aber gerade in der Spitzengruppe lag, lässt sich vermuten, dass die Studenten höhere Testergebnisse erzielen als die Studentinnen.

(F4) Gibt es einen Unterschied im durchschnittlichen Testergebnis der Studienanfängerinnen im Vergleich mit den Studienanfängern?

Im Bezug auf die verschiedenen Schuldauern, welche die Studienanfängerinnen und -anfänger durchlaufen haben, zeigten sich in den in Abschnitt 3.4 vorgestellten Untersuchungen verschiedene Effekte auf die mathematischen Fähigkeiten. In allen Fällen lagen dabei kleine Effekte vor und in einigen Fällen nur in Interaktionen. Daher werden hier keinen großen Unterschiede erwartet werden, auch die Richtung des Effekts in unklar.

(F5) Unterscheiden sich die durchschnittlichen Testergebnisse in Abhängigkeit von der Schuldauer?

(F5b) Gibt es dabei Unterschiede für die verschiedenen Bundesländer?

Für die Zeit zwischen dem Verlassen der Schule und dem Start des Studiums zeigten sich in den Untersuchungen aus Abschnitt 3.5 kleine Vorteile für Studierende, welche früher das Studium aufgenommen haben, im Vergleich mit den verzögert beginnenden Studierenden. Entsprechend lässt sich vermuten, dass die Studierenden mit weniger Verzögerung die besseren Testergebnisse erzielen. Hierbei stellt sich insbesondere die Frage, ob es für die verschiedenen mathematischen Gebiete Unterschiede gibt, die von der zeitlichen Lage im Curriculum abhängen könnten. Ebenso stellt sich die Frage, wo Schülerinnen und Schüler am Ende der Schulzeit im Verhältnis zu den Studienanfängerinnen und -anfänger einzuordnen sind.

(F6) Unterscheiden sich die durchschnittlichen Testergebnisse der Studienanfängerinnen und -anfänger in Abhängigkeit von der Dauer zwischen Schulabschluss und Studienaufnahme für alle drei definierten Gruppen?

(F6b) Wo kann das durchschnittliche Testergebnis der Schülerinnen und Schüler im Verhältnis der drei Gruppen von Studierenden eingeordnet werden?

(F6c) Gibt es dabei Unterschiede für verschiedene mathematische Inhaltsbereiche?

Für die Kurswahl ergaben alle dargestellten Studien aus Abschnitt 3.6 große Vorteile für die Schülerinnen und Schüler, welche einen Kurs auf erhöhtem Anforderungsniveau bzw. Ähnliches gewählt haben. Entsprechend kann vermutet werden, dass sich dieser Unterschied auch hier zeigt.

(F7) Gibt es Unterschiede in den durchschnittlichen Testergebnissen in Abhängigkeit vom gewählten Kurstyp an der Schule?

Nachdem die sieben bisherigen Fragen einzelne Aspekte isoliert betrachten, stellt sich die Frage, welche der Einflüsse sich bei genseitiger Kontrolle als wesentlich erweisen, und, welche darüber hinaus die größten Einflüsse besitzen. Entsprechend der vorgestellten Studien aus Abschnitt 3 kann die Kursform als der Faktor mit dem höchsten Einfluss erwartet werden.

(F8) Welche der Variablen tragen unter gegenseitiger Kontrolle zur Erklärung der Testergebnisse bei?

(F8b) Wie groß ist dabei die erklärte Varianz der Variablen?

Bezogen auf die in der Schule gezeigten Leistungen zeigten sich in den vorgestellten Untersuchungen mittlere Korreleationen zu davon unabhängigen mathematischen Tests. Dabei stellten sich insbesondere die Korrelationen für Schülerinnen und Schüler in Kursen mit erhöhtem Anforderungsniveau als größer heraus als bei einem Besuch eines Kurses auf grundlegendem Anforderungsniveau. Entsprechend können auch hier mittlere bis hohe Korrelationen erwartet werden.

(F9) Wie groß sind die Korrelationen zwischen Testergebnis sowie Abiturdurchschnittsnote und letzter Mathenote?

Werden nun wiederum die bisher einzeln betrachteten Schulleistungen gemeinsam mit allen weiteren Variablen zur Klärung der Varianzen des Testergebnisses verwendet, so lässt sich erwarten, dass diese den größten Beitrag leisten. Es ist unklar, ob die weiteren Variablen zusätzliche Beiträge liefern werden.

(F10) Wie groß ist der Anteil der Varianz in den Testergebnisse, welcher durch die schulischen Leistungsdaten erklärt wird?

(F10b) Wie groß ist der Erklärungsbeitrag der weiteren erhobenen Variablen über die Verwendung der schulischen Leistungsdaten hinaus?

Neben dem Gesamttestergebnis werden auch die einzelnen Bereiche separat ausgewertet. Dabei stellt sich zunächst die Frage, wie stark der Zusammenhang zwischen den verschiedenen Bereichen ist. Hier sind mittlere bis hohe Korrelationen erwartbar. Weiterhin soll untersucht werden, ob die erhobenen Prädiktorvariablen bei gegenseitiger Kontrolle unterschiedliche Einflussstärken auf die Ergebnisse für die mathematischen Teilgebiete besitzen.

(F11) Wie stark ist der Zusammenhang zwischen den acht erhobenen mathematischen Teilgebieten?

(F11b) Gibt es Unterschiede in den Einflüssen auf die Testergebnisse in den Teilgebieten zwischen den erhobenen Variablen?

## 4.2 Klausurergebnisse

Für den Zusammenhang der Klausurergebnisse des ersten Semesters mit den erhobenen Variablen werden das Bestehen sowie die Zensuren betrachtet. Dazu werden zunächst das Antreten einer Klausur, dann das Bestehen der Klausur sowie zuletzt die Zensur in der Klausur untersucht. Die letzten beiden Analysen werden dabei einzeln für die beiden Vorlesungen "Differential- und Integralrechnung 1" und "Analytische Geometrie und lineare Algebra 1" durchgeführt. Für das Bestehen der Klausur soll als Erstes die Frage untersucht werden, ob es Unterschiede in der Bestehensquote zwischen den Teilnehmenden im Propädeutikum und den Nicht-Teilnehmenden gibt. Für die Interpretationen ist dabei nicht völlig klar, ob Unterschiede durch Wirkungen oder Auswahleffekte, in unklare Richtung, oder beides erzeugt werden.

(K1) Gibt es zwischen den Teilnehmenden des mathematischen Propädeutikums und den Nicht-Teilnehmenden Unterschiede in den Bestehensquoten der Klausuren am Ende des ersten Semesters?

Nun, bezogen auf die Teilnehmenden des Propädeutikums wird untersucht, ob es in den Testleistungen Unterschiede gibt zwischen den Studierenden, welche am Ende des ersten Semester an einer Modulprüfung teilgenommen haben, und den Studierenden, die an keiner Modulprüfung teilgenommen haben. Dabei ist davon auszugehen, dass die Studierenden, welche zu einer Modulprüfung antreten, die besseren Testleistungen erbracht haben.

(K2) Gib es einen Unterschied in den Testergebnissen der Studierenden, die am Ende des ersten Semesters eine Modulprüfung angetreten haben, haben zu den Studierenden, die keine Modulprüfung angetreten haben?

Darauf aufbauend wird untersucht, bezogen auf die beiden Veranstaltungen einzeln, ob sich das Bestehen oder Nicht-Bestehen mit Hilfe der drei leistungsbezogenen Variablen, die vor dem Semester erhoben wurden, vorhersagen lässt. Aufgrund der vorgestellten früheren Untersuchungen aus Abschnitt 3.8 lassen sich die Abiturnote oder die Testleistung als bester Einzelprädiktor vorhersehen. Eine Verbesserung der Prädiktion durch die weiteren Variablen oder die einzelnen Testgebiete lässt sich nur auf kleiner Ebene vermuten. Dabei soll hier insbesondere darauf verzichtet werden, die erhobenen Variablen jeweils einzeln zu untersuchen. Wie Halverscheid und Pustelnik (2013) für einen früheren Jahrgang gezeigt hat, sind hier für Geschlecht, Schuldauer sowie verzögerte Schulaufnahme höchstens kleine Effekte zu vermuten. Insbesondere ist davon auszugehen, dass die weiter zurückliegenden Eigenschaften einen nur geringen Einfluss auf Klausurergebnisse besitzen.

(K3a) Wie gut lässt sich das Bestehen in den Klausuren am Ende des ersten Semesters mit Hilfe der Leistungsbezogenen Variablen "Testergebnis","Abiturnote" und "letzte Mathematiknote" voraussagen?

(K3b) Welche dieser drei Variablen ist der beste Prädiktor?

(K3c) Welchen zusätzlichen Beitrag leisten die weiteren erhobenen Variablen?

(K3d) Lässt sich die Prädiktion verbessern durch Verwendung der acht einzelnen Testbereiche?

Nach dem Bestehen der einzelnen Klausuren stellt sich auch die Frage, ob die Ergebnisse in den Modulen des ersten Semesters vorhersagbar sind. Hierbei lässt sich wiederum von der Abiturnote oder der Testleistung als bester Einzelprädiktor ausgehen. Eine Verbesserung mit Hilfe der weiteren Variablen ist dabei möglich. Auch hier ist nicht davon auszugehen, dass sich die Prädiktion durch Verwendung der einzelnen Testbereiche deutlich verbessert.

(K4a) Wie gut können die Klausurnoten in den Klausuren am Ende des ersten Semesters mit Hilfe der leistungsbezogenen Variablen "Testergebnis" ,"Abiturnote" und "letzte Mathematiknote" voraussagen?

(K4b) Welche dieser drei Variablen ist der beste Prädiktor?

(K4c) Welchen zusätzlichen Beitrag leisten die weiteren erhobenen Variablen?

(K4d) Lässt sich die Prädiktion verbessern durch Verwendung der acht einzelnen Testbereiche?

Es stellt sich jedoch die Frage, ob der Zusammenhang zwischen Klausurergebnissen höher ist für Testbereiche, die spezifisch zu den Vorlesungen passen. Alternativ dazu ist die Frage, ob die in der Schule später liegenden Bereiche eine bessere Prädiktion besitzen, unabhängig vom spezifischen Inhalt.

(K5a) Lassen sich zwischen den inhaltlich passenden Testbereichen höhere Korrelationen zu Klausurergebnisse finden?

(K5b) Hängt die Prädiktion der einzelnen Testgebiete von der Stellung in der Schule ab?

Insgesamt geht es in allen Fragen also um die leistungsbezogenen Voraussetzungen der Studienanfängerinnen und -anfänger sowie die Klausurergebnissen. Ein besonderes Interesse der Untersuchung gilt dabei der Unterscheidung verschiedener mathematischer Inhaltsbereiche. Im Bezug auf die Voraussetzungen bedeutet dies die Unterteilung in acht mathematische Bereiche, die auch in der späteren Prädiktion Verwendung finden. Für die Leistungen im ersten Semester werden entsprechend die Modulprüfungen in beiden Veranstaltungen betrachtet und verglichen.

# 5 Methodik

In diesem Kapitel sollen die methodischen Grundlagen der Arbeit vorgestellt werden. Dazu wird als Erstes der verwendete Test in seiner Konstruktion, seinen Inhalten und testtheoretischen Modellierung betrachtet. Anschließend werden die Erhebung der Stichprobe beschrieben sowie alle erhobenen Hintergrundvariablen der Testteilnehmenden vorgestellt. Als Letztes werden kurz die Methoden der Auswertung dargestellt, aufgeteilt nach der Auswertung für das Testergebnis und für die Klausurergebnisse nach einem Semester.

## 5.1 Testkonstruktion

Der verwendete Test wird vor dem Beginn des Studiums eingesetzt. Daher wurden für die Konstruktion nur mathematische Inhalte verwendet, welche die Studienanfängerinnen und -anfänger in der Schule behandelt haben sollen. Gleichzeitig sollen auch nur die Inhalte behandelt werden, die für das Studium im ersten Semester explizit relevant sind. Dies ergibt für die inhaltliche Testkonstruktion auf der einen Seite den Blick nach hinten in die Schulzeit und die dort erlernten Inhalte und auf der anderen Seite den Blick nach vorne auf die Vorlesungen im ersten Semester und die dort nützlichen Inhalte.
Als Grundlage für das aus der Schule erworbene Wissen dienen die Bildungsstandards (Kultusministerkonferenz, 2004, 2012) bzw. davon abgeleitet die Kerncurricula für Niedersachsen (Niedersächsisches Kultusministerium, 2006, 2009). Dabei werden aufgrund der Zielgruppe, Studierende in Mathematik, Physik und Informatik, die Kurse auf erhöhtem Anforderungsniveau sowie eine umfassende Kenntnis dieser Inhalte erwartet. Es ist selbstverständlich nicht davon auszugehen, dass tatsächlich alle Studierenden einen Kurs auf erhöhtem Anforderungsniveau besucht haben (siehe dazu auch Abschnitt 6.7). Dennoch entsprechen diese Kenntnisse den Voraussetzungen für ein Studium in Mathematik bzw. einem Studienfach mit hohem mathematischen Anteil. Ebenso stammen nicht alle Studienanfängerinnen und -anfänger aus Niedersachsen und wurden daher nach anderen Vorgaben unterrichtet. Dabei sind aufgrund der bundesweiten Vorgaben durch die Einheitlichen Prüfungsanforderungen (Kultusministerkonferenz, 2002) Bildungsstandards die mathematischen Inhalte vereinheitlicht worden, sodass keine großen Abweichungen zu erwarten sind, eine Durchsicht der entsprechenden Dokumente bestätigt dies für die Bundesländer mit wesentlichem Studierendenanteil. Die neuen überarbeiteten Curricula für die niedersächsischen Gymnasien spielen für die erhobene keine Rolle, da diese erst später Gültigkeit erhalten haben.
Für die Inhalte der Vorlesungen im ersten Semester "Differential- und Integralrechnung 1" sowie "Analytische Geometrie und lineare Algebra 1" die-

nen zwei Vorlesungsmanuskripte von Dozierenden der Universität Göttingen (Witt, 2012; Kersten, 2005). Dabei ist auch hier nicht die Erwartung, dass für alle Jahre diese Manuskripte verwendet werden. Dennoch stellen sie eine Grundlage für die Vorlesungsinhalte dar, welche zwar von Jahr zu Jahr variieren, aber für die Frage der hilfreichen Schulinhalte weniger unterschiedlich zu erwarten sind.
Nun werden auf Grundlage dieser beiden Vorlagen die Überschnitte identifiziert. Dabei stellt das Kerncurriculum mit den inhaltsbezogenen Kompetenzen die Ausgangsbasis dar. Von den dort formulierten inhaltsbezogenen Kompetenzen ausgehend wird untersucht, ob diese für die Vorlesungen hilfreich oder notwendig sind. Dabei werden die Kompetenzen nicht jeweils einzeln für sich betrachtet, sondern zusammengehörig zu größeren Einheiten zusammengefasst. Für diese zusammengehörigen Felder werden dann Items formuliert. Dabei dienen die im Kerncurriculum als Kompetenzen definierten Inhalte als Vorgaben der mathematischen Bereiche.
Für die mathematischen Gebiete, welche für die Universität von Bedeutung sind, ergeben sich dabei zwei Stränge aus dem Kerncurriculum, welche durch die beiden Vorlesungen des ersten Semesters gegeben sind. Der erste Strang bezieht sich auf die Vorlesung "Differential- und Integralrechnung 1". Dieser folgt der inhaltsbezogenen Kompetenz "Funktionaler Zusammenhang". Diese Kompetenz beginnt in Kerncurriculum mit dem Einführen von Funktionen und ihren Darstellungen bis zur Klasse 6. Dabei werden zunächst lineare Funktionen betrachtet. Hier werden dann in der Jahrgangsstufe 8 die quadratischen Funktionen hinzugefügt. Bis zum Ende der Klasse 10 werden dann entsprechend Potenzfunktionen sowie die trigonometrischen Funktionen eingeführt und im Sinne von Wachstumsverhalten interpretiert. Die Sekundarstufe II entwickelt dann die Differentialrechnung und anschließend die Integralrechnung.
Im Bezug auf die Vorlesung "Analytische Geometrie und lineare Algebra" stellt die Leitidee "Zahlen und Operationen" die zentrale inhaltbezogene Kompetenz des Kerncurriculums dar. Diese beginnt bis zur Klasse 6 mit dem Rechnen mit Bruchzahlen. Die Jahrgangsstufen 7 und 8 stellen dann den Übergang zu den reellen Zahlen und die Einführung von Variablen und Termen dar. Hier werden dann insbesondere auch Gleichungen und Ungleichungen betrachtet. Bis zur Jahrgangsstufe 10 kommen dann Gleichungen mit Potenzen hinzu. In der Sekundarstufe II wird dann die Vektorgeometrie behandelt.
In beiden Vorlesungen kommen natürlich auch aus den anderen Bereichen beschriebene Inhalte hinzu, ebenso aus den prozessbezogenen Kompetenzen. Dabei stellen insbesondere Stochastik und elementare Geometrie der Ebene keine Gebiete der Vorlesungen des ersten Semesters dar. Die acht sich ergebenden Bereiche werden im nächsten Abschnitt genauer beschrieben.
Bei der Auswahl der Aufgaben wurde dabei eher auf eine breite Abdeckung der Gebiete geachtet als spezifische Aspekte tief zu behandeln. Dazu beste-

hen einige der Items aus dem Aufrufen von Faktenwissen (siehe dazu das Beispiel zur Differentialrechnung im nächsten Abschnitt), welches nur aufgerufen werden muss. Daneben lassen sich die meisten Aufgaben mit kurzschrittigen Verfahren lösen, die bereits bekannt und eingeübt sind (siehe dazu das Beispiel zu den Grundlagen der Algebra im nächsten Abschnitt). Dabei ist der rechnerische Aspekt bewusst klein gehalten. Diese beiden Itemarten lassen sich damit dem prozeduralen Wissen zuordnen (Hiebert & Lefevre, 1986). Konzeptuelles Wissen wird im Wesentlichen in Verbindung mit Darstellungswechseln von Funktionen direkt abgefragt.

## 5.2 Inhaltliche Beschreibung der Testkonstruktion

Im Folgenden soll nun ein kurzer Überblick über die mathematischen Inhalte des Test gegeben werden, welche sich aus der oben beschriebenen Konstruktion ergeben. Dazu wird für jeden der acht Testbereiche der Inhalt kurz zusammengefasst und anschließend ein Beispielitem angegeben zur Konkretisierung angegeben. Am Ende werden die Bereiche jeweils mit Hilfe der niedersächsischen Kerncurricula in eine Doppelstufe eingeordnet.

### 5.2.1 Grundlagen der Algebra

Im ersten Testabschnitt werden vor allem die Themen Äquivalenz, Zahlenräume, Bruchrechnung und lineare Gleichungen behandelt. Im Folgenden ist ein Beispielitem[3] zur Äquivalenzumformung von Gleichungen angegeben.

> *Beispielitem:*
> Gegeben ist die Gleichung $y = x + 5$. Welche der folgenden Aussagen gelten für alle Lösungen $(x, y)$ mit $x, y \in \mathbb{R}$ ?
> ☐ $y + 2 = x + 7$
> ☐ $13 \cdot x = 13 \cdot y - 65$
> ☐ $y > 5$

Der Übergang zu den Bruchzahlen stellt Inhalte der Jahrgangsstufe 5 und 6 dar. Lineare Gleichungen sowie der Umgang mit Termen und Äquivalenz fallen in die siebte Klasse. Der Umgang mit linearen Funktionen aus Klasse 8 wird hier nicht behandelt, jedoch wird die Kenntnis von reellen Zahlen vorausgesetzt. Die Inhalte lassen sich somit im Wesentlichen den Jahrgangsstufen 6 und 7 zuzuordnen.

[3] Die genaue Beschreibung der Antwortformate findet im nächsten Abschnitt statt. Für die folgenden acht Beispielitems ergibt sich aus der Fragestellung, dass entweder genau eine der Antworten richtig ist oder alle richtigen Aussagen ausgewählt werden müssen. Im Beispielitems für den Bereich "Polynome" muss eine freie Antwort gegeben werden und ist entsprechend gekennzeichnet.

### 5.2.2 Gleichungen und Ungleichungen

Der Abschnitt zu Gleichungen und Ungleichungen behandelt quadratische Gleichungen, Systeme linearer Gleichungen und lineare Ungleichungen sowie Betragsungleichungen.

> *Beispielitem:*
> Welche ist die Lösungsmenge $\mathbb{L}$ der Ungleichung $(x-2)^2 > 4$?
> ☐ $\mathbb{L} = \{x \in \mathbb{R} | x < 4\}$
> ☐ $\mathbb{L} = \{x \in \mathbb{R} | x < 4\}$
> ☐ $\mathbb{L} = \{x \in \mathbb{R} | x < 0\}$
> ☐ $\mathbb{L} = \{x \in \mathbb{R} | x > 0\}$
> ☐ $\mathbb{L} = \{x \in \mathbb{R} | x > 0$ und x $< 4\}$
> ☐ $\mathbb{L} = \{x \in \mathbb{R} | x > 4$ oder x $< 0\}$

Das Lösen von linearen und quadratischen Gleichungen wird im Kerncurriculum bis zum Ende der Jahrgangsstufe 8 im Rahmen der Kompetenz "Zahl und Operation" behandelt. Dabei fällt die Einführung linearer Gleichungen und auch Ungleichungen in die Jahrgangsstufe 7, wird dann in der achten Klasse in Bezug auf Funktionen und damit auch grafisch vertieft. In dieser Jahrgangsstufe werden auch quadratische Gleichungen und Funktionen ausführlich untersucht.

### 5.2.3 Polynome

Im Abschnitt zu Polynomen werden graphisch Umkehrfunktionen, das Erkennen des Graphen von Polynomfunktionen sowie deren Symmetrie, Grenzverhalten und Nullstellenanzahl behandelt.

> *Beispielitem:*
> Wie viele reelle Nullstellen besitzen die folgenden Polynome?
> ■ $f_1(x) = x^4 + x^2$ besitzt genau ___ verschiedene reelle Nullstellen?
> ■ $f_2(x) = x^4 - x^2$ besitzt genau ___ verschiedene reelle Nullstellen?

Die Themen dieses Abschnitts werden damit über verschiedene Jahrgangsstufen behandelt, werden aber alle unter der Kompetenz "Funktionaler Zusammenhang" gefasst. Die verschiedenen Eigenschaften der Polynome fällt im Wesentlichen unter die Funktionsuntersuchung ganzrationaler Funktionen in der Jahrgangsstufe 10, auch wenn Ansätze dazu in vorigen Jahrgangsstufen aufgebaut werden, wie auch das Beispielitem, welches im Rahmen von quadratischen Gleichungen behandelt werden könnte.

### 5.2.4 Exponential- und Logarithmusfunktionen

Der Bereich beinhaltet Items zum Berechnen von Potenzen und Logarithmen, die zugehörigen Rechenregeln sowie die Graphen von Exponential- und Logarithmusfunktionen mit den zugehörigen Grenzwerten sowie Definitions- und Wertebereichen.

> *Beispielitem:*
> Welche der folgenden Aussagen sind wahr?
> ☐ Die Funktion $f(x) = e^x$ ist für alle reellen Zahlen definiert.
> ☐ Die Funktion $f(x) = \log_e(x)$ ist für alle reellen Zahlen definiert.
> ☐ Die Funktion $f(x) = e^x$ besitzt die gesamten reellen Zahlen als Wertebereich.
> ☐ Die Funktion $f(x) = \log_e(x)$ besitzt die gesamten reellen Zahlen als Wertebereich.

Nachdem bis Jahrgangsstufe 6 Potenzen mit natürlichen Exponenten untersucht werden, werden diese im Rahmen der Kompetenz "Zahl und Operation" bis Klasse 10 auf rationale Exponenten erweitert. Bis zu dieser Jahrgangsstufe werden ebenso Potenz- und Exponentialfunktionen untersucht. Auch wenn der Logarithmus im Kerncurriculum nicht explizit erwähnt wird, so sollen Schülerinnen und Schüler doch "Gleichungen in einfachen Fällen algebraisch mit Hilfe von Umkehroperationen lösen"(Niedersächsisches Kultusministerium, 2006, p. 27). Dabei findet die Einführung von Potenzfunktionen und -gleichungen in der Jahrgangsstufe 9 statt. Eine Verwendung im Rahmen verschiedener Wachstumsvorgänge findet dann in der Sekundarstufe II statt, wird hier aber nicht weiter betrachtet.

### 5.2.5 Trigonometrie

Der Bereich Trigonometrie behandelt die Umrechnung von Grad- und Bogenmaß, die Definition trigonometrischer Funktionen am Einheitskreis sowie die Eigenschaften der trigonometrischen Funktionen wie Nullstellen und Extrema sowie deren Umkehrbarkeit.

> *Beispielitem:*
> Welche der folgenden Aussagen sind wahr?
> ☐ Die Sinusfunktion besitzt an der Stelle $4\pi = 720°$ ein lokales Extremum.
> ☐ Die Nullstellenmenge des Sinus ist gegeben durch alle ganzzahligen Vielfachen von $\pi = 180°$.
> ☐ Die Kosinusfunktion besitzt an der Stelle $\pi = 180°$ ein lokales Minimum.
> ☐ Die Nullstellenmenge der Kosinusfunktion ist gegeben durch $N = \{l \cdot \pi + \frac{\pi}{2} | l \in \mathbb{Z}\}$.

Die geometrische Einführung der trigonometrischen Funktionen findet sich im Kerncurriculum bis zum Ende der Jahrgangsstufe 10 unter der Kompetenz "Größe und Messen". Diese Inhalte werden entsprechend in der neunten Klasse behandelt. Die Behandlung als Funktionen wird ebenso bis zum Ende der Klasse 10 unter "Funktionaler Zusammenhang" beschrieben und findet in der 10 Klasse statt. Die funktionale Betrachtung steht dabei hier im Fokus.

### 5.2.6 Vektorrechnung

Der Abschnitt zur Vektorrechnung beinhaltet Lagebeziehungen und Abstände von Punkten, Geraden und Ebenen im dreidimensionalen Raum. Ebenso werden Addition und Multiplikation von Vektoren und Matrizen behandelt.

> *Beispielitem:*
> Liegen die folgenden Punkte auf einer gemeinsamen Geraden im $\mathbb{R}^3$?
> $P(1/1/1)$ $Q(2/3/4)$ $R(2/3/5)$
> ☐ ja
> ☐ nein

Der Abschnitt bezieht sich auf die Leitideen "Messen" sowie "Räumliches Strukturieren/ Koordinatisieren" aus der Sekundarstufe II.

### 5.2.7 Differentialrechnung

Der Bereich beinhaltet die Berechnung der Ableitungen bekannter Funktionen, einfache Rechenregeln sowie das graphische Ableiten von Funktionen und die Untersuchung von Extrem- und Wendestellen.

> *Beispielitem:*
> Kreuzen Sie die 1. Ableitung der Funktion $g(x) = \mathrm{e}^x$ auf ihrem maximalen Definitionsbereich an. Hier bezeichnet e die Eulersche Zahl.
> ☐ $g\prime(x) = x \cdot \mathrm{e}^{x-1}$
> ☐ $g\prime(x) = \mathrm{e}$
> ☐ $g\prime(x) = \log(x)$
> ☐ $g\prime(x) = \mathrm{e}^{x^2}$
> ☐ $g\prime(x) = \mathrm{e}^x$
> ☐ $g\prime(x) = x \cdot \mathrm{e}^x$

Die Einführung findet unter der Kompetenz "Funktionaler Zusammenhang" bis zum Ende der Jahrgangsstufe 10 statt. Hierbei werden die Definition erarbeitet und auf ganzrationale Funktionen angewendet sowie einfache Regeln betrachtet. Die Betrachtung weiterer Funktionen sowie ausführliche Funktionsuntersuchungen, welche Steigungs- und Krümmungsverhalten enthalten, findet dann in der Sekundarstufe II statt.

### 5.2.8 Integralrechnung

Der Abschnitt zur Integralrechnung behandelt die Kenntnis der Stammfunktionen der bekannten Funktionen, die einfachen Integralregeln, das Berechnen von Flächeninhalten, inklusive Orientierung sowie das graphische Finden einer Stammfunktion.

> *Beispielitem:*
> Kreuzen Sie eine Stammfunktion $G$ der Funktion $g(x) = \sin(x)$ auf ihrem maximalen Definitionsbereich an.
> ☐ $G(x) = \cos(x)$
> ☐ $G(x) = -\cos(x)$
> ☐ $G(x) = \frac{1}{\tan(x)}$
> ☐ $G(x) = \tan(x)$
> ☐ $G(x) = -\sin(x)$
> ☐ $G(x) = \sin(x)$

Der gesamte Bereich der Integralrechnung lässt sich dem Bereich "Funktionaler Zusammenhang" in der Sekundarstufe II zuordnen.
Eine Zusammenfassung der Einordnung der Testbereiche in die Jahrgangsstufen zeigt Tabelle 1. Hier werden ebenso die in einigen zukünftigen Tabellen aus Platzgründen verwendeten Abkürzungen aufgeführt.

| Testbereich | Abkürzung | zugeordnete Jahrgangsstufen |
|---|---|---|
| Grundlagen der Algebra | Alg | Jahrgangsstufen 6/7 |
| Gleichungen und Ungleichungen | Glei | Jahrgangsstufen 7/8 |
| Polynome | Poly | Jahrgangsstufen 10 |
| Exponential- und Logarithmusfunktionen | Exp | Jahrgangsstufen 9/10 |
| Trigonometrie | Trig | Jahrgangsstufen 10 |
| Vektorrechnung | Vec | Jahrgangsstufen 11 |
| Differenzialrechnung | Diff | Jahrgangsstufen 10/11 |
| Integralrechnung | Int | Jahrgangsstufen 11 |

Tabelle 1: Zuordnung der Testbereiche zu den Schuljahrgängen

## 5.3 Testmodellierung

In diesem Abschnitt werden nun die testtheoretischen Eigenschaften des Tests bzw. der Items vorgestellt. Dies beginnt mit der Konstruktion und der anschließenden Modellierung durch das 1-dimensionale Rasch-Modell. Anschließend werden Kriterien für die Modellgültigkeit überprüft. Dazu gehört insbesondere die Betrachtung der acht einzelnen Testbereiche.
Alle formulierten Items werden mit "richtig" oder "falsch" bewertet, es gibt

also insbesondere keine Teilpunkte. Dabei liegen, wie die Beispiele des vorigen Abschnitts bereits zeigen, drei verschiedene Antwortformate vor: Multiple choice, single choice und ein freies Antwortfeld. Dabei bedeuten freie Antworten in jedem Fall die Angabe einer Zahl oder eines einfachen Ausdrucke, insbesondere werden keine Begründungen oder längeren Angaben erwartet. Die Unterscheidung zwischen single choice und multiple choice ist dabei für die Teilnehmenden deutlich erkennbar durch die Verwendung verschiedener Kästchenformen für das Ankreuzen.
Für die Modellierung des Test wird nun das eindimensionale Raschmodell (Rasch, 1960) verwendet. Dazu werden aus den empirisch beobachteten Daten die Itemschwierigkeiten sowie die Personenfähigkeit der Personen bestimmt. Dabei werden zunächst alle Items als Gesamttest modelliert und anschließend die acht Teilgebiete jeweils einzeln. Die Modellgültigkeit wird dabei sowohl für die einzelnen Bereiche als auch für den Gesamttest überprüft, sodass Items ausgeschlossen werden, wenn sie in einem der beiden Tests nicht aufgenommen werden. Der Gesamttest stellt also die Vereinigung der acht einzelnen Testbereiche dar.
Die Modelltestung geschieht dabei an den Daten der Propädeutikumsteilnehmenden des Jahrgangs 2011, die nicht Teil der Hauptuntersuchung sind. Dabei werden entsprechend die Voraussetzungen des Rasch-Modells jeweils global als Gesamttest und lokal auf Item- bzw. Pesonenebene getestet. Ausschluss von problematischen Items liefert dann eine Verkürzung der verwendeten Testitems auf eine modellkonforme Anzahl. Diese Items wurden dann für die hier untersuchten Stichproben verwendet. In einem letzten Schritt wird dann das eindimensionale Rasch-Modell verglichen mit verallgemeinerten Modellen. Dazu werden als Modellgütekriterien verschiedene Informationsindizes verwendet.
Als Modelltests wurden entsprechend der Empfehlung aus Strobl (2015) zunächst global der Andersen-Test durchgeführt. Dabei wurde jeweils der Median-Split als Teilungskriterium verwendet. Dieser Test untersucht die Subgruppeninvarianz. Die globale Überprüfung der Itemhonogenität findet mit Hilfe des Martin-Löf-Tests statt, wobei wiederum ein Median-Split vorgenommen wird. Bei Nichtgültigkeit des Raschmodells wurden jeweils Items auf Grundlage des itemspezifischen Wald-Tests entfernt bis zur Modellgültigkeit entfernt. Ein Ausschluss von Personen wurde nicht durchgeführt
Ein Vergleich des Rasch-Modells mit dem allgemeineren Birnbaummodells bzw. Mixed-Raschmodellen mit zwei und drei Subgruppen zeigte, dass das Raschmodell die beste Passung auf die Daten ergab. Hier wird entsprechend der Empfehlung von Rost (1996, p. 329) für dünn besetzte Antwortmuster der BIC-Index verwendet. Die Ergebnisse der Tests sind im Anhang in den Tabellen 31 bis 33 zu finden.
Insgesamt ergaben sich damit der Gesamttest sowie die acht Teiltests als Rasch-konform modellierbar. Die verkürzten Versionen wurden dann für die Erhebungen verwendet, wobei zwischen den Jahrgängen keine Veränderun-

gen in den Items vorgenommen wurden.
Die Betrachtung der acht Bereiche durch die Testkonstruktion zeigt also, dass sich der Test als ein 1-dimensionales Konstrukt verstehen lässt, sodass hier die einzelnen Bereiche so testtheoretisch nicht bestätigt werden. Da die Bereiche allerdings aufgrund ihrer Konstruktion verschiedene mathematische Bereiche definieren, sollen diese im Folgenden auch als einzelne Skalen interpretiert werden. Dies wird auch durch die Korrelationen in Tabelle 2 bestätigt.

| | Alg | Glei | Poly | Exp | Trig | Vec | Diff | Int |
|---|---|---|---|---|---|---|---|---|
| Alg | | | | | | | | |
| Glei | 0,380 | | | | | | | |
| Poly | 0,374 | 0,525 | | | | | | |
| Exp | 0,362 | 0,509 | 0,482 | | | | | |
| Trig | 0,390 | 0,570 | 0,558 | 0,516 | | | | |
| Vec | 0,375 | 0,516 | 0,535 | 0,441 | 0,554 | | | |
| Diff | 0,356 | 0,494 | 0,508 | 0,454 | 0,552 | 0,562 | | |
| Int | 0,303 | 0,532 | 0,533 | 0,468 | 0,592 | 0,568 | 0,573 | |

Tabelle 2: Korrelationen zwischen Teilgebieten

Es zeigen sich also ausschließlich Korrelationen, die einen Wert unter 0,6 besitzen. Diese Korrelationen sind als Korrelationen zwischen verschiedenen mathematischen Gebieten als moderat zu bezeichnen. Die Höhe dieser Korrelationen spricht damit auch für die Betrachtung und Auswertung der einzelnen Teiltestgebiete.

## 5.4 Stichprobenziehung

Die Stichprobe wurde über einen Zeitraum von vier Jahren erhoben, von 2013 bis 2016. Die Studierenden der Stichprobe sind Teilnehmende am mathematischen Propädeutikum der Universität Göttingen. Das Propädeutikum stellt einen Brückenkurs für Studierende dar, welcher sich vor allem an die Studierenden der Mathematik, der Physik, der Informatik sowie die Lehramtsstudierenden in Mathematik richtet und auf die Vorlesungen des ersten Semesters dieser Studiengänge abgestimmt ist.
Dabei ist die Teilnahme für alle Studierende, unabhängig des Studiengangs, offen. Das mathematische Propädeutikum ist eine reines Präsenzangebot, welches kostenlos und freiwillig ist, wenn die Teilnahme auch dringend empfohlen wird. Das Propädeutikum beginnt acht Wochen vor Studienbeginn und dauert drei Wochen.
Die Erhebung fand jeweils am ersten Tag des Propädeutikum statt, im Anschluss an eine kurze Begrüßung. Die Teilnahme an der Testung war freiwillig. Allerdings erfolgt aufgrund der Bearbeitung eine Rückmeldung an die Teilnehmenden mit einer Empfehlung für den Besuch von spezifischen

Workshops des Propädeutikums, sodass die Testdurchführung so auch einen vorgesehenen Teil der Veranstaltung darstellt. Dies führt zu einer umfassenden Teilnahme an dem Test. Für Teilnehmende, die erst zu einem späteren Tag mit dem Besuch der Veranstaltung begannen, wurden innerhalb der jeweils ersten Woche Nachholtermine angeboten, um eine möglichst vollständige Stichprobenziehung der Propädeutikumsteilnehmenden zu gewährleisten. Durchgeführt und beaufsichtigt wurde der Test durch studentische Hilfskräfte, welche auch Übungsgruppen für das Propädeutikum gegeben haben. Diese waren angewiesen keine inhaltlichen Fragen beantworten, und insbesondere auch keine Definitionen für unbekannte Begriffe geben. Auch die Zeitvorgabe von 90 Minuten sollte klar eingehalten werden. Eine Ankündigung vor der Veranstaltung oder Vorbereitungszeit zwischen Begrüßung und Testdurchführung gab es für die Teilnehmenden nicht, sodass ein explizites Lernen nicht erwartet wurde.
Als eine Vergleichsgruppe zu den Studierenden wurden außerdem Schülerinnen und Schüler mit dem selben Test untersucht. Die Schülerinnen und Schüler, welche den Test durchgeführt haben, stammen aus Schulen der Umgebung Göttingens. Es wurden jeweils die Kurse mit erhöhtem Anforderungsniveau angesprochen, welche in ihrem letzten Schuljahr waren, also die zwölfte Klasse besuchten. Die Teilnahme am Test war freiwillig und fand außerhalb der Unterrichtszeit in den Gebäuden der Schule statt. Als Anreiz wurde den Schülerinnen und Schülern eine Rückmeldung zu ihren Testergebnissen gegeben.
Eine Genehmigung des Datenschutzbeauftragten für die Verwendung der erhobenen Daten in Verknüpfung mit den Klausueergebnisse liegt vor[4]. Ebenso die Zustimmung der Auswertungen der Schülerinnen und Schüler.

### 5.5 Darstellung der Variablen

Für die Auswertungen werden die Einflüsse von sechs nominalen bzw. ordinalen Variablen auf die Testleistungen untersucht. Zusätzlich werden zwei als metrisch angesehene unabhängige Variablen, die Abiturdurchschnittsnote, sowie die letzte Mathenote, betrachtet. Die Testleistung ist dabei auch metrisch skaliert. Als weitere Kontrollvariable für die Stichprobenbildung wird das Jahr der Datenerhebung für die Studierenden in die Auswertungen aufgenommen: Damit ergeben sich die folgenden unabhängigen Variablen mit ihren jeweiligen Ausprägungen:

- Studiengang
  - Mathematik (1-Fach-Bachelor)
  - Physik (1-Fach-Bachelor)
  - Informatik (1-Fach-Bachelor)

[4] Die Genehmigung wurde am 02.12.2014 erteilt.

  - Gymnasiallehramt Mathematik (2-Fach-Bachelor Profil Lehramt)
- Geschlecht
  - männlich
  - weiblich
- Schuldauer
  - 12 Jahre (G8)
  - 13 Jahre (G9)
- Verzögerter Studienbeginn
  - direkte Studienaufnahme
  - Studienaufnahme genau im Jahr nach Schulabschluss
  - Studienaufnahme mindestens zwei Jahre nach Schulabschluss
- Bundesland
  - Niedersachsen
  - Hessen
  - Nordrhein-Westfalen
  - Schleswig-Holstein
- Kursform
  - Kurs auf erhöhtem Anforderungsniveau (oder Ähnliches)
  - Kurs auf grundlegendem Anforderungsniveau (oder Ähnliches)
- Abiturdurchschnittsnote
  - Notenskala von 1,0 bis 4,0 in Schritten von 0,1
- letzte Mathenote
  - Punkteskala von 15 Punkten bis 0 Punkten in Schritten von 1 Punkt
- Jahr des Studienbeginns
  - 2013
  - 2014
  - 2015
  - 2016

Für den Vergleich mit den Schülerinnen und Schülern, für die keine Hintergrundvariablen vorliegen, werden diese in zwei Variablen als eigene Gruppe betrachtet. Im Vergleich der Studiengänge werden sie als fünfter Studiengang betrachtet; dabei wird auch die globalere Unterscheidung Studierende gegenüber Schülerinnen und Schüler gemacht. Im Vergleich des verzögerten Studienbeginns werden die Schülerinnen und Schüler ebenso als eine vierte Gruppe in den Analysen verwendet.
Als abhängige Variable der Auswertungen wird zunächst das Testergebnis betrachtet, welches ein metrisches Skalenniveau besitzt. Ebenso werden die acht Testergebnisse in den einzelnen mathematischen Gebieten verwendet. In den Analysen zu Ergebnissen der Klausurleistungen in den Universitätsveranstaltungen werden Klausurergebnisse als metrische Variablen auf der Notenskala von 1 bis 5 mit den üblichen Schritten von 0,33 verwendet.
Die Klausurnoten werden für die beiden Veranstaltungen "Differential- und Integralrechnung 1" sowie "Analytische Geometrie und lineare Algebra 1" untersucht für die Anfangsjahre 2013 und 2014. Im Jahr 2013 war diese Veranstaltung eine Pflichtveranstaltung für die Studierenden in Mathematik, in Physik sowie für die Lehramtsstudierenden. Ab dem Jahr 2014 galt dies nicht mehr für die Studierenden in Physik.
Für jede Veranstaltung gibt es jeweils zwei angebotene Klausurtermine im Anschluss an das erste Semester, die hier untersucht werden.[5] Dabei werden für die Studierenden jeweils das Ergebnis der ersten Klausur und das Ergebnis der letzten geschriebenen Klausur als zwei Variablen unterschieden. Als erste Klausur werden dabei nur die Studierenden betrachtet, welche die erste Klausur in ihrem ersten Studiensemester geschrieben haben. Der Jahrgang wird entsprechend auch hier für die Analysen verwendet. Für alle Klausurbetrachtungen wird weiterhin die dichotome Variable "bestanden/nicht bestanden" abgeleitet.

## 5.6 Statistische Auswertungsverfahren

Die Auswertung untergliedert sich in zwei Abschnitte. Zunächst werden die Einflüsse der oben beschriebenen Variablen auf die Ergebnisse im Test untersucht, anschließend wird untersucht, welchen Einfluss die Variablen und zusätzlich das Testergebnis auf Klausurergebnisse nach dem ersten Semester besitzen.
Für die Auswertungen der Testergebnisse werden die Mittelwerte der verschiedenen Gruppen mit Hilfe von Varianzanalysen (ANOVA) und linearen Regressionen (Backhaus, Erichson, Plinke & Weiber, 2015) untersucht. Im Fall von nur zwei Ausprägungen der unabhängigen Variablen geht diese in einen $t$-Test über und die Ergebnisse werden in diesem Rahmen berichtet. Die Auswertungen der Klausuren orientiert sich an einer Untersuchung zu

[5] Die weitere Möglichkeit, die "Differential- und Integralrechnung 1" über das Sommerstudium zu besuchen, wird nicht berücksichtigt.

Klausurergebnissen in Physik (Sorge et al., 2016).
Varianzanalysen besitzen als Voraussetzung eine Normalverteilung der Messwerte in den untersuchten Gruppen. Allerdings haben sich Varianzanalysen als sehr robust gegenüber, auch starken, Verletzungen (Schmider, Ziegler, Danay, Beyer & Bühner, 2010) erwiesen. In Bezug auf die verwendeten Teiltests ist die Normalverteilungsannahme der Varianzanalyse nicht immer gegeben, diese werden allerdings dennoch aufgrund der Robustheit dennoch durchgeführt. Die Verwendung von nichtparametrischen Alternativen ginge mit einem Verlust an Teststärke einher.
Die zweite Voraussetzung der Varianzhomogenität wird mit Hilfe des Levene-Test überprüft. Da bei ungleich besetzten Gruppengrößen die Varianzanalyse nicht robust gegenüber dieser Verletzung ist, wird bei einer Verletzung dieser Annahmen als Alternative zur regulären ANOVA ein Welch-Test (Welch, 1947) berechnet, welcher die Varianzhomogenität nicht als Voraussetzung besitzt[6].
Für die Post-hoc-Gruppenvergleiche bei einem signifikanten Einfluss eines Prädiktors werde für die paarweisen Gruppenvergleiche Korrekturen der Fehler 1. Art nach Hochberg (Benjamini & Hochberg, 1995) durchgeführt. Diese Korrekturen ergeben eine höhere Teststärke als oft verwendete Korrekturen nach Bonferroni, die als sehr konservativ einzuschätzen sind. Entsprechend werden bei signifikanter Varianzheterogenität Post-Hoc-Tests nach Games-Howell verwendet (Ruxton & Beauchamp, 2008).
Für die gemeinsame Betrachtung von Variableneinflüssen werden gemeinsame mehrfaktorielle Varianzanalysen für die gegenseitige Kontrolle der verschiedenen Einflussfaktoren durchgeführt. Dabei werden in Analysen mit allen Einflussvariablen nur Haupteffekte betrachtet. In Fällen mit nur zwei der Einflussvariablen hingegen werden auch die Interaktionseffekte erster Ordnung betrachtet.
Als Maß der Effektstärke der einzelnen Varianzanalysen wird global $\eta^2$ verwendet (Cohen, 1973), welches sich als erklärter Varianzanteil interpretieren lässt. Für den Einfluss der verschiedenen Variablen in einer gemeinsamen Analyse werden partielle $\eta^2$ angegeben (Levine & Hullett, 2002).
Für die Post-hoc-Gruppenvergleiche dient Cohens $d$ als Maß der Effektstärke und der Einfluss wird nach Cohens Klassifikation (Cohen, 1992) berichtet. Aufgrund der unbalancierten Gruppengrößen wird als Schätzer für die Varianz die gepoolte Varianz (Hartung, Knapp & Sinha, 2011) verwendet. Für die Einschätzungen der Effektstärke ist zu beachten, dass aus empirischer Sicht diese Klassifikation sehr streng ist (Gignac & Szodorai, 2016; Hemphill, 2003) und in wenigen Fällen außerhalb von Experimentaldesigns hohe Effektstärken festgestellt werden entsprechend werden die Effektstärken in

[6]Da bei Verwendung der Welch-Tests nicht-ganzzahlige Freiheitsgrade auftreten, werden bei Verteilungen mit zwei Parametern diese durchgängig durch Semikola getrennt, und so von den Dezimalzahlen unterschieden.

der Diskussion eingeordnet.
Werden auch metrische Variablen als Prädiktoren in den Analysen verwendet, werden zunächst die bivariaten Korrelationen berechnet. Anschließend werden unter Verwendung, sowohl der metrischen als auch der nicht metrischen, Variablen gleichzeitig, multiple lineare Regressionen (Backhaus et al., 2015) durchgeführt. Dabei werden die nichtmetrischen Variablen jeweils durch einzelne Dummy-Variablen ersetzt. Entsprechend der Reihenfolge in Abschnitt 5.5 werden die jeweils ersten Kategorien als Referenzkategorien verwendet. Die Aufnahme der Prädiktorvariablen in das Regressionsmodell geschieht dabei schrittweise.
Das Bestimmtheitsmaß $R^2$ als Verallgemeinerung des Korrelationsquadrats für die linearen Regressionen dient der Beschreibung der Modellgüten der linearen Regressionen und lässt sich wiederum als aufgeklärte Varianz des Modells interpretieren. Auch hier werden Stärken nach Cohen berichtet.
Für die Betrachtung der dichotomen Variable des Klausurbestehens werden logistische Regressionen durchgeführt. Auch hier werden die Prädiktorvariablen schrittweise hinzugefügt. Für die Güte werden hier Mc Faddens $R^2$ berichtet (Menard, 2000), sowie die Anzahl der korrekt zugeordneten Fälle durch das Regressionsmodell. Für den Vergleich nicht-geschachtelter Modelle dient weiterhin das Informationskriterium AIC (Akaike, 1998). Die Bedeutung der einzelnen Prädiktorvariablen wird aufgrund der unstandardisierten Koeffizienten berichtet.
Im Umgang mit fehlenden Werten werden jeweils die Personen ausgeschlossen, sodass die einzelnen Untersuchungen sich jeweils auf verschiedene Gruppengrößen beziehen.

# 6 Deskriptive Beschreibung der Stichprobe

Die gesamte Stichprobe besteht aus $N = 1014$ Teilnehmenden, davon sind $N_{Stud} = 890$ Studienanfängerinnen und -anfänger sowie $N_{SuS} = 124$ Schülerinnen und Schüler. Die Studierenden verteilen sich auf vier Jahrgänge, während die Schülerinnen und Schüler zu einem Zeitpunkt erhoben wurden und von acht verschiedenen Schulen jeweils aus einem Kurs mit erhöhtem Anforderungsniveau stammen. Zu den Schülerinnen und Schülern wurden keine weiteren Hintergrundvariablen erhoben. Im Folgenden werden nun die erhobenen Daten für die Studierenden im Bezug auf die verschiedenen Variablen, welche als unabhängige Variablen für die Analysen verwendet werden, beschrieben. Als Erstes werden der Anteil der Stichproben an allen Studierenden der untersuchten Studiengänge sowie die Verteilung auf die Jahrgänge dargestellt.

## 6.1 Einordnung in die Gesamtanfängerzahlen

Im Folgenden soll die erhobene Stichprobe im Propädeutikum in das Verhältnis zu allen Studienanfängerinnen und -anfängern gesetzt werden. Die Anzahl der Anfängerinnen und Anfänger sowie der Anteil der Stichprobe für jeden Studiengang sind in Tabelle 3 dargestellt (Georg-August-Universität Göttingen, 2018).

| Anfangsjahr | Mathematik | Physik | ang. Informatik | LA Mathe |
|---|---|---|---|---|
| 2013 | 69 | 186 | 75 | 72 |
| | 39,1 % | 53,2 % | 38,7 % | 45,8 % |
| 2014 | 84 | 195 | 102 | 110 |
| | 35,7% | 50,3 % | 34,3 % | 28,2 % |
| 2015 | 115 | 193 | 164 | 107 |
| | 33,0% | 49,7 % | 34,1 % | 35,5 % |
| 2016 | 126 | 158 | 165 | 94 |
| | 46,0% | 43,7 % | 21,2 % | 33,0 % |
| $\Sigma$ | 394 | 732 | 506 | 383 |
| | 38,8 % | 49,5 % | 30,6 % | 34,7 % |

Tabelle 3: Anteil der Propädeutikumsteilnehmenden an allen Studienanfängerinnen und -anfängern nach Studiengang

Insgesamt erfasst die erhobene Stichprobe ungefähr 30 % bis 50 % aller Studienanfängerinnen und Anfänger. Dieser Anteil unterscheidet sich zwischen den vier Studiengängen deutlich. Ebenso ergeben sich zwischen den Erhebungsjahren deutliche Unterschiede für die einzelnen Studiengänge. Dabei ergibt sich bei den Schwankungen über die Jahre kein einheitliches Bild zwischen den Studiengängen.

## 6.2 Studiengang

Für die Auswertung werden vier Studiengänge unterschieden: 1-Fach-Bachelor Mathematik, 1-Fach-Bachelor Physik, 1-Fach-Bachelor Angewandte Informatik sowie die Lehramtsstudierenden in Mathematik mit beliebigem zweiten Fach. Alle weiteren Studiengänge werden nicht einzeln ausgewertet, da sie nicht der Zielgruppe des Propädeutikums und der Mathematikvorlesungen entsprechen und auch eine zu geringe Anzahl für statistische Auswertungen aufweisen.

| Anfangsjahr | Mathematik | Physik | ang. Informatik | LA Mathe | gesamt |
|---|---|---|---|---|---|
| 2013 | 27 | 99 | 29 | 33 | 200 |
| | 13,5 % | 49,5 % | 14,5 % | 16,5 % | |
| 2014 | 30 | 98 | 35 | 31 | 214 |
| | 14,0 % | 45,8 % | 16,4 % | 14,5 % | |
| 2015 | 38 | 96 | 56 | 38 | 267 |
| | 14,2 % | 36,0 % | 21,0 % | 14,2 % | |
| 2016 | 58 | 69 | 35 | 31 | 209 |
| | 27,8 % | 33,0 % | 16,8 % | 14,8 % | |
| zusammen | 153 | 362 | 155 | 133 | 890 |
| | 17,2 % | 40,7 % | 17,4 % | 14,9 % | |

Tabelle 4: Studiengang der Studierenden nach Anfangsjahr (Prozentangaben beziehen sich jeweils auf die Studierenden des Jahrgangs)

Wie in Tabelle 4 zu sehen ist, verteilen sich die Studierenden, mit Ausnahme des Jahres 2015, gleichmäßig auf die Erhebungsjahre. Im Vergleich der verschiedenen Studiengänge zeigt sich, dass die Verteilung auf die Studiengänge ziemlich konstant ist. Nur im Jahr 2016 ist der Anteil der Studierenden in Mathematik verdoppelt im Gegensatz zu den vorigen Jahrgängen, ebenso verringert sich zwischen den Jahren 2014 und 2015 der Anteil der Physikstudierenden an der Gesamtteilnehmendenanzahl. Hier erhöht sich gleichzeitig der Anteil der Studierenden in angewandter Informatik leicht.
Insgesamt stellen die Studierenden in Physik mit 40 % den größten Anteil dar, während die drei weiteren Studiengänge jeweils einen ungefähr gleich großen Anteil besitzen. Von der gesamten Stichprobe fallen ungefähr 90 % der Studierenden in einen der vier untersuchten Studiengänge.

## 6.3 Herkunftsbundesland

Von den 890 Studierenden stammt der erwartungsgemäß größte Anteil mit 55,5 % aus Niedersachsen, dies entspricht 494 Studierenden. Darüber hinaus gibt es drei weitere Bundesländer, aus denen mehr als 5 % der Studienanfängerinnen und -anfänger der Stichprobe stammen: Hessen mit 88

Studierenden, dies entspricht 9,9 % der Stichprobe; 78 Studierenden aus Nordrhein-Westfalen, was 8,8 % entspricht; sowie Schleswig-Holstein mit 53 Studierenden und damit einem Anteil von 6,0 % der Stichprobe. Studierende, die ihre Hochschulzugangsberechtigung im Ausland erworben haben, machen mit 23 Studierenden einen Anteil von 2,3 % der Stichprobe aus, werden aber nicht gesondert ausgewertet.

## 6.4 Dauer des Schulbesuchs

Der Anteil der Studierenden mit einem Abitur nach 13 Jahren beträgt insgesamt 28,2 %. Dabei zeigt sich dieser nach einem Abfall zwischen den Anfangsjahren 2013 und 2014 als relativ konstant. Dieser Abfall fällt zusammen mit den Umstellungen auf die zwölfjährige Abiturdauer in Hessen und Nordrhein-Westfalen zum Abiturjahrgang 2013. Die Umstellung in Schleswig-Holstein im Jahr 2016 ist hier nicht zu erkennen.

| Anfangsjahr | G8 | G9 | Σ |
|---|---|---|---|
| 2013 | 117 | 75 | 192 |
| | 60,9 % | 39,1 % | |
| 2014 | 158 | 51 | 209 |
| | 75,6 % | 24,4 % | |
| 2015 | 181 | 72 | 253 |
| | 71,5 % | 28,5 % | |
| 2016 | 158 | 43 | 201 |
| | 78,6 % | 21,4 % | |
| Σ | 614 | 241 | 855 |
| | 71,8 % | 28,2 % | |

Tabelle 5: Schuldauer der Studierenden nach Anfangsjahr (Prozentangaben beziehen sich jeweils auf die Studierenden des Jahrgangs, welche eine Angabe zur Schuldauer gemacht haben)[7]

Eine Auffälligkeit ergibt sich im Verhältnis zwischen der Schuldauer und der verzögerten Studienaufnahme. Studierende mit einem verzögerten Studienbeginn von mindestens zwei Jahren besitzen zu 56,9 % eine Schuldauer von 13 Jahren, also wesentlich häufiger als in der gesamten Stichprobe. Dies kann wesentlich dadurch erklärt werden, dass Studierende, die mit einer längeren Verzögerung an die Universität kommen, in einem Zeitraum ihr Abitur gemacht haben, in dem 13 Jahre Schulzeit in fast allen Bundesländern die Regel war.

[7] Durch das Runden der Prozentangaben addieren sich in dieser und den folgenden Tabellen die Werte nicht immer genau zu 100 %.

## 6.5 Verzögerter Studienbeginn

Durch die Differenz aus dem Jahr des Schulabschlusses und dem Jahr des Studienbeginns wird die Verzögerung der Studienaufnahme errechnet. Hierbei ergeben sich Verzögerungen von 0 Jahren bis zu 27 Jahren. Alle Studierenden mit einer Verzögerung von mehr als einem Jahr werden dabei zusammengefasst, sodass sich die drei Kategorien "keine Verzögerung", "genau 1 Jahr Verzögerung" und "mindestens 2 Jahre Verzögerung" ergeben.

| Studiengang | 0 Jahre | 1 Jahr | mind. 2 Jahre | Σ |
|---|---|---|---|---|
| Mathematik | 105 | 31 | 13 | 149 |
| | 70,5 % | 20,8 % | 8,7 % | |
| Physik | 267 | 71 | 19 | 357 |
| | 74,8 % | 19,9 % | 5,3 % | |
| ang. Informatik | 94 | 21 | 38 | 153 |
| | 61,4 % | 13,7 % | 24,8 % | |
| LA Mathe | 79 | 35 | 19 | 133 |
| | 59,4 % | 26,3 % | 14,3 % | |
| Σ | 545 | 158 | 89 | 792 |
| | 68,8 % | 19,9 % | 11,2 % | |

Tabelle 6: Verzögerung des Studienbeginns nach Studiengang (Prozentangaben beziehen sich jeweils auf die Studierenden des Studiengangs, welche eine Angabe zur Verzögerung gemacht haben)

Wie in Tabelle 6 zu sehen ist, haben ungefähr zwei Drittel aller Studierenden ihr Studium ohne eine Verzögerung aufgenommen. Dabei liegt der Anteil in den Studiengängen Mathematik und Physik noch etwas höher, während die Studierenden in angewandter Informatik und im Lehramt Mathematik zu ca. 60 % ihr Studium im Jahr des Schulabschlusses aufnehmen. Auffällig ist weiterhin für die Studierenden in angewandter Informatik ein hoher Anteil an Studierenden mit mindestens zwei Jahren verzögerter Studienaufnahme unter den Studierenden, welche nicht direkt ihr Studium beginnen.

Im bundesweiten Vergleich fällt die Quote der verzögerungslosen Studienaufnahme im Jahr 2012 mit 35 % wesentlich geringer aus als für alle vier Studiengänge der hier untersuchten Stichprobe (Middendorff et al., 2013). Dabei zeigt sich ein Trend über die letzten Jahre hin zu einer geringeren Verzögerungsquote. So haben im Jahr 2003 nur 25 % der Studierenden ihr Studium ohne Verzögerung aufgenommen.

Bei den Verzögerungen ergibt sich dabei ein deutlicher Unterschied in Abhängigkeit des Geschlechts. Im Jahr 2003 lag der Anteil der Frauen mit direkter Studienaufnahme mit 35 % deutlich über dem für Männer mit 16 %. Bis zum Jahr 2012 verringerte sich der Abstand auf Quoten von 40 % und 31 % bei Frauen bzw. Männer. Die Differenz von 10 % mehr direkten Studienaufnahmen von Frauen finden sich auch in der Stichprobe wieder, wenn auch auf

höherem Niveau.
Dabei ist zu beachten, dass die Abschaffung der Wehrpflicht im Jahr 2011 sich noch auf beide dieser Angaben auswirkt[8]. Als Median der verzögerten Studienaufnahme ergibt sich für 2012 ein Wert von 14 Monaten, dies entspricht der Verzögerung um 1 Jahr.

## 6.6 Geschlecht

Von den 890 Studierenden sind insgesamt 635, also 71,3 % männlich, und 255 weiblich, dies entspricht einem Anteil von 28,7 % der Stichprobe. Dabei ergibt sich eine unterschiedliche Verteilung auf die verschiedenen Studiengänge, wie in Tabelle 7 zu sehen ist.

| Studiengang | männlich | weiblich | Σ |
|---|---|---|---|
| Mathematik | 102 | 51 | 153 |
| | 66,7 % | 33,3 % | |
| Physik | 279 | 83 | 362 |
| | 77,1 % | 22,9 % | |
| ang. Informatik | 134 | 21 | 155 |
| | 86,5 % | 13,5 % | |
| LA Mathe | 58 | 75 | 133 |
| | 43,6 % | 56,4 % | |
| Σ | 573 | 230 | 803 |
| | 71,4 % | 28,6 % | |

Tabelle 7: Geschlechterverteilung nach Studiengang (Prozentangaben beziehen sich jeweils auf die Studierenden des Studiengangs, welche eine Angabe zum Geschlecht gemacht haben)

Während also für den Mathematik-Lehramtsstudiengang der Anteil von Frauen knapp über der Hälfte liegt, ist der Frauenanteil in den drei anderen Studiengängen jeweils deutlich geringer als der Anteil der Männer. Dabei unterscheidet er sich auch zwischen diesen drei Studiengängen noch einmal erheblich: So liegt der Frauenanteil unter den Mathematikstudierenden mit einem Drittel höher als im Physikstudiengang mit ungefähr einem Fünftel und in angewandter Informatik mit einem Anteil von einem Achtel der Studierenden.

## 6.7 Schulleistungsdaten

Als Leistungsdaten aus der Schule werden die Abiturdurchschnittsnote sowie die letzte Mathematiknote der Studierenden verwendet. Die letzte Mathe-

[8] Die aktuellere Sozialerhebung 2016 (Middendorff et al., 2017) weist die entsprechende Zahl nicht gesondert aus.

matiknote stellt also die Note in einer Abiturprüfung dar oder die Mathematiknote des letzten Semesters, falls keine Abiturprüfung in Mathematik abgelegt wurde. Andere Arten der Hochschulzugangsberechtigung werden nicht berücksichtigt. Außerdem wurde erhoben, ob die Studierenden einen Kurs mit grundlegendem Anforderungsniveau (gA-Kurs) oder mit erhöhtem Anforderungsniveau (eA-Kurs), bzw. einen entsprechenden Kurs in anderen Bundesländern, besucht haben.
In Tabelle 8 sind die entsprechenden Ergebnisse zu sehen, angegeben in Abhängigkeit der verschiedenen Studiengängen. Dabei werden die beiden Zensuren, insbesondere auch in Mathematik nicht, nach der Art des besuchten Kurses nicht unterschieden angegeben. Hier fallen die Unterschiede in den Durchschnittsnoten nur extrem gering aus. Ebenso sind in Tabelle 8 als Vergleich die durchschnittlichen Ergebnisse der Abiturprüfungen in Niedersachsen aus dem Jahr 2016 zu sehen (Niedersächsisches Kultusministerium, 2016a).

| Studiengang | Abiturnote | letzte Mathenote | Anteil eA-Kurs |
|---|---|---|---|
| Mathe | 1,8 | 13,4 | 84,2 % |
| Physik | 1,8 | 12,6 | 79,7 % |
| ang. Informatik | 2,5 | 10,0 | 57,7 % |
| LA Mathe | 2,2 | 11,6 | 72,3 % |
| ∅ | 2,0 | 12,1 | 75,1 % |
| Niedersachsen | 2,6 | 7,6 (eA)<br>7,4 (gA - schriftlich)<br>8,4 (gA - mündlich) | |

Tabelle 8: Abiturdurchschnittsnote, letzte Mathenote und Kursniveau nach Studiengang (Prozentangaben beziehen sich jeweils auf die Studierenden des Studiengangs, welche eine Angabe zum Kursniveau gemacht haben); sowie Prüfungsergebnisse der niedersächsischen Abiturprüfungen 2016

Die Abiturdurchschnittsnoten der vier Gruppen liegen jeweils über der Abiturdurchschnittsnote in Niedersachsen für das Jahr 2016, die bei 2,58 lag. Wie Abschnitt 6.3 zeigt, kommt der Großteil der Studierenden aus Niedersachsen, sodass die niedersächsischen Schulnotenschnitte als Vergleich verwendet werden. Auch wenn dies der höchste Abiturschnitt im Vergleich der Bundesländer ist (Statista, o. J.), liegen die Studierenden in Mathematik und in Physik immer noch unter dem Abiturschnitt aller deutschen Bundesländer. Damit besitzen die Studienanfängerinnen und -anfänger in Mathematik sowie Physik also deutlich überdurchschnittliche Abiturdurchschnittsnoten. Auch die Lehramtsstudierenden besitzen einen Abiturdurchschnitt, der deutlich besser als der niedersächsische Abiturschnitt ist, während der Abiturschnitt der Informatikstudierenden im Durchschnitt Niedersachsens liegt. Damit zeigen sich bei insgesamt guten Schulleistungen noch wesentliche Un-

terschiede zwischen den verschiedenen Studiengängen.
Die Durchschnittspunktzahl der Abiturprüfungen in Mathematik für Niedersachsen im Jahr 2016 liegt im Kurs mit erhöhtem Anforderungsniveau bei 7,56 Punkten (Niedersächsisches Kultusministerium, 2016b) und im Kurs auf grundlegendem Niveau schriftlich bei 7,37 Punkten sowie mündlich bei 8,44 Punkten[9]. Im Vergleich dazu zeigen die vier Studierendengruppen wesentlich höhere Durchschnittspunktzahlen. Die Mathematikstudierenden erbringen mit nahezu 6 Notenpunkten mehr als der niedersächsische Durchschnitt die besten Leistungen und besitzen hier ebenso einen leicht besseren Durchschnitt als die Studierenden in Physik. Auch die Lehramtsstudierenden erbringen eine weit überdurchschnittliche Leistung, welche 4 Punkte über dem Durchschnitt Niedersachsens liegt. In Bezug auf die letzte Mathematiknote besitzen die Studierenden in angewandter Informatik wiederum den geringsten Durchschnittswert, dieser liegt jedoch immer noch 2,5 Punkte höher als der Durchschnitt der Klausur im Kurs mit erhöhtem Anforderungsniveau.
In der letzten Mathematiknote zeigen sich für alle vier Studiengänge bessere Durchschnittswerte als im Vergleich mit dem Durchschnitt Niedersachsen. Der Abstand liegt dabei zwischen einer und zwei ganzen Noten, sodass dieser als beträchtlich zu bezeichnen ist. Dabei ergeben sich dennoch auf hohem Niveau auch zwischen den verschiedenen Studiengängen noch Differenzen von bis zu 3 Notenpunkten.
Die Quoten des eA-Kurs-Besuchs liegen entsprechend der Reihung in den Durchschnittsnoten zwischen knapp 58 % für Informatikstudierenden und knapp 85 % unter den Mathematikstudierenden. Dabei liegen die Studierenden im Lehramt zwischen den beiden vorher genannten Gruppen leicht unter der Gruppe der Studierenden in Physik.

[9]Dabei ist zu beachten, dass die Benotung der zentralen Abiturklausur im eA-Kurs nachträglich angehoben wurde. (Niedersächsisches Kultusministerium, 2016c)

# 7 Auswertung der Testergebnisse

In diesem Abschnitt wird untersucht, welchen Einfluss die erhobenen Variablen auf das Ergebnis des Tests besitzen. Zunächst werden dazu im Folgenden die Einflüsse der Faktoren einzeln ausgewertet. Dafür werden für die sieben vorhandenen nichtmetrischen Variablen einfaktorielle Varianzanalysen durchgeführt, in Einzelfällen auch zweidimensional. Anschließend werden in einer gemeinsamen Analyse die verschiedenen Faktoren und ihr gemeinsamer Einfluss auf die Testleistung miteinander verglichen.
Als Nächstes werden dann die Variablen, welche Schulleistungen ausdrücken, betrachtet, zunächst einzeln, und anschließend gemeinsam mit den vorher betrachteten Variablen. Dabei werden wiederum die verschiedenen Einflüsse auf ihre Stärken hin untersucht.
Am Ende wird dargestellt, ob es Unterschiede zwischen den verschiedenen mathematischen Testgebieten gibt, indem die Einflüsse der Variablen auf die verschiedenen Gebiete miteinander verglichen werden. Dies geschieht insbesondere für die Kursform und die Verzögerung des Studienbeginns, für die sich hier Unterschiede ergeben. In allen Untersuchungen werden für den verzögerten Studienbeginn sowie den Studiengang die Schülerinnen und Schüler als Vergleichsgruppe betrachtet.

## 7.1 Erhebungszeitpunkt

Da die Stichprobe aus Studienanfängerinnen und -anfängern vier verschiedener Anfangsjahre des Studiums als Erhebungszeitpunkte besteht, wird zunächst ein möglicher Einfluss des Erhebungszeitpunktes auf die Ergebnisse im Gesamttest untersucht.

| Anfangsjahr | Anzahl | Mittelwert | Standardabweichung |
|---|---|---|---|
| 2013 | 200 | 0,90 | 0,86 |
| 2014 | 214 | 1,07 | 0,94 |
| 2015 | 267 | 1,02 | 0,97 |
| 2016 | 209 | 1,12 | 0,95 |
| Σ | 890 | 1,03 | 0,93 |

Tabelle 9: Mittelwert und Standardabweichung der Testergebnisse in Abhängigkeit des Anfangsjahrs

Die Varianzanalyse ergibt für die vier Erhebungszeitpunkte mit $F(3; 886) = 2,00; p = 0,112$ keinen signifikanten Einfluss des Erhebungszeitpunkts auf das Testergebnis. Die zugehörigen Mittelwerte sind in Tabelle 9 zu sehen. Auch paarweise Vergleiche der Erhebungszeitpunkte ohne Fehleradjustierungen ergeben für keinen der Vergleiche einen signifikanten Unterschied.

## 7.2 Studiengang

Die Untersuchung des Einflusses des Studiengang erfolgt in zwei Schritten: Als Erstes werden die Studienanfängerinnen und -anfänger, unabhängig von ihrem Fach, mit den Schülerinnen und Schülern verglichen. Anschließend werden die vier Studiengänge untersucht. Dabei werden die Schülerinnen und Schüler im Nachhinein als ein eigener fünfter Studiengang zum Vergleich eingeordnet.
Für die Mittelwerte der $N_{Stud} = 890$ Studienanfängerinnen und -anfänger und der $N_{SuS} = 124$ Schülerinnen und Schülern ergibt sich mit $t(199,78) = 0,01; p = 0,936$ keine signifikante Differenz zwischen den Gruppen. Da in diesem Vergleich keine Varianzgleichheit vorliegt, wird, wie in Abschnitt 5.6 beschrieben, der angepasste Welch-Test durchgeführt.
Der Einfluss des Studiengangs wird nun mit Hilfe einer Varianzanalyse mit den vier Ausprägungen betrachtet. Diese zeigt einen signifikanten Einfluss des Studiengangs, $F(3;799) = 78,66; p \leq 0,001; \eta^2 = 0,23$, und beschreibt einen starken Effekt auf das Testergebnis.

| Studiengang | Anzahl | Mittelwert | Standardabweichung |
|---|---|---|---|
| Mathe | 153 | 1,53 | 0,86 |
| Physik | 362 | 1,37 | 0,81 |
| Informatik | 155 | 0,45 | 0,79 |
| LA Mathe | 133 | 0,60 | 0,75 |
| SuS | 124 | 1,02 | 0,66 |
| Σ | 927 | 1,09 | 0,89 |

Tabelle 10: Mittelwert und Standardabweichung der Testergebnisse in Abhängigkeit des Studiengangs, inklusive Schülerinnen und Schülern

Post-hoc-Gruppenvergleiche ergeben, dass die Unterschiede der Studierenden in Mathematik und Physik auf der einen Seite zu den Studierenden in Informatik und des Lehramts auf der anderen Seite jeweils mit $p \leq 0,001$ signifikant sind, während die Unterschiede innerhalb der beiden Paare keine Signifikanz aufweisen. Die zugehörigen Mittelwerte und Standardabweichungen sind in Tabelle 10 zu sehen.
Für den Unterschied zwischen den Studienanfängerinnen und -anfängern in Mathematik und den Studienanfängerinnen und -anfängern in Informatik liegt mit $d = 1,31$ ebenso ein starker Effekt vor wie im Vergleich der Mathematikstudierenden mit den Studierenden des Lehramts, dessen Stärke $d = 0,96$ beträgt. Die entsprechenden Vergleich der Physikstudierenden ergeben ebenso einen starken Effekt mit $d = 1,14$ gegenüber den Studierenden in Informatik sowie einen mittleren Effekt von $d = 0,78$ gegenüber den Lehramtsstudierenden.
Werden nun die Schülerinnen und Schüler für die gleiche Analyse als fünfter Studiengang betrachtet, so ergibt sich auch in diesem Fall, mit $F(4;374,06) =$

$60,11; p \leq 0,001$, ein signifikanter Einfluss des Studiengangs. Dabei liegt der Mittelwert der Schülerinnen und Schüler im Testergebnis zwischen den beiden gefundenen Gruppen von Studierenden. Es ergeben sich mit $d = 0,65$ zu den Studierenden in Mathematik ein mittlerer Effekt und mit $d = 0,45$ zu den Studierenden in Physik ein kleiner Effekt zu Ungunsten der Schülerinnen und Schüler. Auf der anderen Seite zeigen die Schülerinnen und Schüler mit Effektstärken von $d = 0,78$ im Vergleich mit den Studierenden in Informatik und $d = 0,38$ im Vergleich mit den Studierenden im Lehramt bessere Testergebnisse von mittlerer bis kleiner Effektstärke.

## 7.3 Herkunftsbundesland

Die gemessene Testergebnisse der Studierenden werden für die vier Herkunftsbundesländer verglichen, deren Gruppengrößen dies ermöglichen (siehe dazu Abschnitt 6.3). Neben den Studierenden aus Niedersachsen werden die Studierenden aus Hessen, Nordrhein-Westfalen und Schleswig-Holstein in die Analyse aufgenommen. Für diese Bundesländer wird wiederum durch eine Varianzanalyse der Einfluss des Bundeslandes auf die Testleistung untersucht.

In Tabelle 11 sind für die vier Bundesländer die Anzahl der Teilnehmerinnen und Teilnehmer sowie ihr Mittelwert und ihre Standardabweichung aufgeführt. Die Varianzanalyse zeigt einen signifikanten Einfluss des Bundeslandes auf das Testergebnis, $F(3;709) = 11,09; p \leq 0,001; \eta^2 = 0,05$, mit kleiner Effektstärke.

| Herkunftsbundesland | Anzahl | Mittelwert | Standardabweichung |
|---|---|---|---|
| Niedersachsen | 494 | 0,89 | 0,92 |
| Hessen | 88 | 1,39 | 0,81 |
| Nordrhein-Westfalen | 78 | 1,19 | 1,03 |
| Schleswig-Holstein | 53 | 1,37 | 0,90 |
| $\Sigma$ | 713 | 1,02 | 0,94 |

Tabelle 11: Mittelwert und Standardabweichung der Testergebnisse in Abhängigkeit des Herkunftsbundeslandes

Die beiden Post-hoc-Vergleiche zwischen Studierenden aus Niedersachsen und Hessen sowie aus Niedersachsen und Schleswig-Holstein zeigen jeweils mit $p \leq 0,001$ bzw. $p = 0,002$ eine signifikante Differenz. Die Effekte entsprechen mit $d = 0,55$ im Vergleich Niedersachsens mit Hessen und $d = 0,52$ im Vergleich mit Schleswig-Holstein jeweils einer mittlere Effektstärke.

Die Differenz zwischen Teilnehmenden aus Niedersachsen und Nordrhein-Westfalen ist mit $p = 0,052$ nicht signifikant. Ebenso zeigen die Post-hoc-Tests, dass für die Studierenden aus den Bundesländern Hessen, Nordrhein-Westfalen und Schleswig-Holstein paarweise keine signifikanten Differenzen vorliegen.

## 7.4 Geschlecht

Zum Vergleich der Studierenden abhängig vom Geschlecht wird ein Welch-Test mit zwei unabhängigen Stichproben durchgeführt, da hier keine Varianzhomogenität zwischen den Gruppen vorliegt.
Die $N_{♂} = 635$ Studienanfänger besitzen mit $\mu_{♂} = 1,09$ ($\sigma_{♂} = 0,97$) einen höheren Mittelwert im Mathematiktest als die $N_{♀} = 255$ Studienanfängerinnen, deren Mittelwert $\mu_{♀} = 0,89$ ($\sigma_{♀} = 0,83$) beträgt. Dies bedeutet eine signifikante Differenz $t(540,79) = 3,12; p = 0,002$ der Leistungen mit einer Effektstärke von $d = 0,21$. Dies beschreibt einen kleinen Effekt.
Um hier die Varianzungleichheit genauer zu untersuchen, sind in Tabelle 12 die Quartilswerte für die beiden Geschlechter getrennt aufgetragen.

| Geschlecht | 1. Quartil | 2. Quartil | 3. Quartil |
|---|---|---|---|
| männlich | 0,36 | 1,07 | 1,80 |
| weiblich | 0,28 | 0,87 | 1,38 |
| gesamt | 0,37 | 1,00 | 1,67 |

Tabelle 12: Quartile der Testergebnisse in Abhängigkeit des Geschlechts

Da die Standardabweichung der Männer größer ist als für die Frauen und diese einen höheren Mittelwert besitzen, zeigt sich entsprechend ein größerer Abstand im oberen Quartil. Dem entsprechend ist die Differenz der Geschlechter im unteren Quartil nur sehr gering. Die Differenz zwischen den Gruppen ergibt sich also im Vergleich der Spitzengruppen.

## 7.5 Dauer des Schulbesuchs

Für den Vergleich der Studierenden mit zwölfjähriger Schulzeit und dreizehnjähriger Schulzeit werden diese beiden Gruppen zunächst unabhängig von ihrem Herkunftsbundesland mithilfe einer einfaktoriellen Varianzanalyse betrachtet. Anschließend wird die Auswertung unter Beachtung einer Interaktion von Schuldauer und Bundesland zweifaktoriell untersucht.
Es befinden sich $N_{12} = 614$ und $N_{13} = 241$ Studienanfängerinnen und Studienanfänger mit einer Schuldauer von zwölf bzw. dreizehn Jahren in der Stichprobe. Dabei liegt der Mittelwert der Studierenden mit einer zwölfjährigen Schulzeit mit $\mu_{12} = 1,21$ ($\sigma_{12} = 0,87$) über dem der Studierenden mit dreizehnjähriger Schulzeit, der $\mu_{13} = 0,69$ ($\sigma_{13} = 0,94$) beträgt. Dieser Unterschied erweist sich mit $t(853) = 7,69; p \leq 0,001$ als signifikant und besitzt mit $d = 0,58$ eine mittlere Effektstärke.
Die Varianzanalyse unter Einbeziehen des Herkunftsbundeslands und zugehörigem Interaktionseffekt zeigt, dass sowohl die beiden Haupteffekte, mit $F(1; 692) = 22,50; p \leq 0,001; \eta_p^2 = 0,031$ für die Schuldauer und mit $F(3, 692) = 23,14; p \leq 0,001; \eta_p^2 = 0,091$ für das Herkunftsbundesland als

auch der Interaktionseffekt mit $F(3; 692) = 6,19; p \leq 0,001; \eta_p^2 = 0,03$ signifikante Einflüsse besitzen. Dabei besitzen das Bundesland einen mittleren und die beiden anderen Einflüsse einen kleinen Effekt.

| Schuldauer | Herkunftsbundesland | Anzahl | Mittelwert | Standard-Abweichung |
|---|---|---|---|---|
| 12 Jahre | Niedersachsen | 381 | 1,11 | 0,86 |
| | Hessen | 52 | 1,51 | 0,77 |
| | Nordrhein-Westfalen | 52 | 1,48 | 0,95 |
| | Schleswig-Holstein | 13 | 1,34 | 1,11 |
| | Σ | 498 | 1,20 | 0,88 |
| 13 Jahre | Niedersachsen | 104 | 0,19 | 0,70 |
| | Hessen | 36 | 1,20 | 0,84 |
| | Nordrhein-Westfalen | 24 | 0,72 | 0,90 |
| | Schleswig-Holstein | 38 | 1,43 | 0,82 |
| | Σ | 202 | 0,67 | 0,93 |
| Σ | | 700 | 1,04 | 0,93 |

Tabelle 13: Mittelwert und Standardabweichung der Testergebnisse in Abhängigkeit der Schuldauer und des Herkunftsbundeslands

In Tabelle 13 sind die zum Interaktionseffekt gehörigen Mittelwerte zu sehen. Es zeigt sich, dass für die Bundesländer Hessen und Schleswig-Holstein der Unterschied der Schuldauer gering ist, im Falle von Schleswig-Holstein auch entgegen der Richtung des Haupteffekts. Auf der anderen Seite ergeben sich für die Studierenden aus Niedersachsen und Nordrhein-Westfalen wesentlich bessere Leistungen im Testergebnis für die Studienanfängerinnen und -anfänger, die eine zwölfjährige Schulzeit besaßen. Dieses unterschiedliche Verhalten innerhalb der Bundesländer macht den Interaktionseffekt aus.

## 7.6 Verzögerung des Studienbeginns

Für die Untersuchung der verzögerten Studienaufnahme wird eine 1-faktorielle Varianzanalyse durchgeführt, welche drei Merkmalsausprägungen für die Studienanfängerinnen und -anfänger besitzt. Anschließend werden als vierte Gruppe, wie in Abschnitt 7.2, die Schülerinnen und Schüler hinzugefügt. Diese stellen hier eine Vergleichsgruppe dar, welche keine Zeit zwischen Abitur und Studienbeginn verbracht hat. Dabei wird in diesem Abschnitt nur der Gesamttest betrachtet, eine Betrachtung der einzelnen Testbereiche findet in Abschnitt 7.10 statt.

Die Varianzanalyse zeigt mit $F(2; 874) = 43,33; p \leq 0,001; \eta^2 = 0,09$ einen signifikanten Einfluss des verzögerten Studienbeginns auf das Testergebnis mit einer mittleren Effektstärke. Für die drei Gruppen der Studierenden ergeben sich die Mittelwerte und Standardabweichungen aus Tabelle 14.

Die Post-hoc-Paarvergleich zeigen paarweise signifikante Differenzen zwischen den drei Gruppen: Dabei ergibt sich mit $p = 0,034$; $d = 0,22$ ein kleiner Effekt zwischen dem direkten Studienbeginn und dem um ein Jahr verzögerten Studienbeginn sowie mit $p \leq 0,001$; $d = 0,95$ ein starker Effekt zwischen dem direkten Studienbeginn verglichen mit dem um mindestens zwei Jahre verzögerten Studienbeginn. Der Vergleich der beiden Gruppen mit verzögerten Studienbeginn ergibt mit $p \leq 0,001$; $d = 0,73$ eine mittlere Effektstärke. Die zugehörigen Mittelwerte zeigt Tabelle 14.

| Verzögerung | Anzahl | Mittelwert | Standardabweichung |
|---|---|---|---|
| keine Verzögerung | 591 | 1,18 | 0,86 |
| genau 1 Jahr | 175 | 0,98 | 0,89 |
| mindestens 2 Jahre | 111 | 0,33 | 0,90 |
| gesamt | 877 | 1,03 | 0,93 |

Tabelle 14: Mittelwerte und Standardabweichungen der Testergebnisse in Abhängigkeit von der Verzögerung des Studienbeginns

Werden nun die Schülerinnen und Schüler mit in die Analyse aufgenommen, sodass vier Gruppen verglichen werden, ergibt sich für die Varianzanalyse mit $F(3; 294,74) = 28,34; p \leq 0,001$ ein ähnliches Ergebnis. In den Gruppenvergleichen mit den $N_{SuS} = 124$ Schülerinnen und Schülern mit einem Mittelwert von $\mu_{SuS} = 1,02$ ($\sigma_{SuS} = 0,66$) zeigt sich nur für die Gruppe von Studierenden mit mindestens zwei Jahren Verzögerung des Studienbeginns eine signifikante Differenz $p \leq 0,001$, welche mit $d = 0,88$ einen starken Effekt anzeigt.
Dabei ist zu bemerken, dass in diesem Fall die Differenz von keinem und genau einem Jahr verzögerten Studienbeginn im Post-hoc-Vergleich keine Signifikanz mehr aufweist, was auf die erhöhte Korrektur durch die Hinzunahme einer weiteren Vergleichsgruppe zurückzuführen ist, sodass die Teststärke reduziert wird.

## 7.7 Zusammenhang mit Schulleistungen

Als Merkmale der Schulleistung wurden die Abiturdurchschnittsnote und die letzte Mathematiknote erfasst, dazu kommt die Frage nach dem Besuch eines Kurses auf erhöhtem Anforderungsniveaus oder auf grundlegendem Anforderungsniveau. Zunächst wird durch einen $t$-Test der Vergleich der verschiedenen Kursniveaus durchgeführt. Anschließend werden die Korrelationen der beiden Schulzensuren mit den Testergebnissen berechnet und nach der belegten Kursform unterschieden. Als letzte Analyse wird eine lineare Regression für das Testergebnis mit Hilfe der drei hier betrachteten Prädiktoren durchgeführt.
Der Vergleich der Studierenden in Abhängigkeit von ihrer Kurswahl ergibt

mit $t(859) = 12,88; p \leq 0,001$ eine signifikante Differenz. Dabei besitzen die $N_{eA} = 621$ Studierenden mit einem Kurs auf erhöhtem Anforderungsniveau, $\mu_{eA} = 1,28$ ($\sigma_{eA} = 0,87$), einen höheren Mittelwert als die $N_{gA} = 240$ Studierenden, die einen Kurs auf grundlegendem Anforderungsniveau, $\mu_{gA} = 0,45$ ($\sigma_{gA} = 0,79$), besuchten. Diese Differenz zeigt dabei mit $d = 0,98$ einen starken Effekt an.
Wie Tabelle 15 zeigt, korrelieren die Testergebnisse mit beiden Leistungsmaßen aus der Schule ähnlich hoch, zwischen den beiden Schulmaßen sind die Korrelationen höher als jeweils zum Testergebnis. Weiterhin zeigt sich, dass alle drei Korrelationen für Studierende mit einem Kurs auf erhöhtem Anforderungsniveau höher ausfallen als für Studierende, die einen Kurs auf grundlegendem Anforderungsniveau besucht haben, insbesondere auch die beiden Schulnoten. Dabei zeigen alle Korrelationen einen starken Effekt an.

| Korrelationen | Test-Abitur | Test-letzte Mathenote | Abitur-letzte Mathenote |
|---|---|---|---|
| erhöhtes Anforderungsniveau | -0,60 (N=612) | 0,61 (N=594) | -0,77 (N=590) |
| grundlegendes Anforderungsniveau | -0,51 (N=232) | 0,51 (N=224) | -0,64 (N=220) |
| alle Teilnehmenden | -0,59 (N=859) | 0,58 (N=834) | -0,73 (N=822) |

Tabelle 15: Korrelationen zwischen den Testergebnisse und den Schulleistungen (Fälle werden paarweise ausgeschlossen, sodass sich pro Zelle verschiedene Studierendenanzahlen ergeben)

Die Regressionsanalyse mit drei unabhängigen Prädiktoren für $N = 810$ klärt mit einem Bestimmtheitsmaß von $R^2 = 0,48$ knapp die Hälfte der Varianz innerhalb der Testergebnisse auf. Damit besitzen die schulischen Leistungsmaße erwartungsgemäß einen hohen Einfluss auf das Testergebnis. Das zugehörige Modell zeigt Tabelle 16. Die Referenzkategorie der Kursform stellt dabei der Kurs mit grundlegendem Anforderungsniveau dar.

| Prädiktor | $B$ | $\beta$ | $t$ | $p$ | Änderung in $R^2$ |
|---|---|---|---|---|---|
| Konstante | -0,304 | | -1,21 | 0,23 | |
| Abiturnote | -0,438 | -0,316 | -8,515 | $\leq 0,001$ | 0,350 |
| Kursform | 0,601 | 0,291 | 11,283 | $\leq 0,001$ | 0,085 |
| letzte Mathenote | 0,101 | 0,310 | 8,391 | $\leq 0,001$ | 0,045 |

Tabelle 16: Lineare Regressionsanalyse mit dem Testergebnis als abhängiger Variable ($B$: unstandardisierte Regressionskoeffizienten, $\beta$: standardisierte Regressionskoeffizienten, $p$: Signifikanzniveau für die Verschiedenheit von 0 des jeweiligen Koeffizienten im Endmodell, Änderung in $R^2$: Vergrößerung des Bestimmtheitsmaß durch schrittweises Einfügen des Prädiktors)

Für die lineare Regression werden nacheinander alle drei Variablen in das Modell aufgenommen und liefern ein signifikanten Beitrag für die Vorhersage. Die Verwendung der Abiturnote als ersten Prädiktor liefert bereits ein Bestimmtheitsmaß von $R^2 = 0,35$. Danach werden dann zusätzlich die Kursform und die letzte Mathematiknote aufgenommen. Das Vorzeichen für die Abiturnote besitzt dabei ein negatives Vorzeichen, da für die Abiturnote kleiner Werte eine bessere Leistung darstellen. Für die letzte Mathematiknote hingegen bedeuten höhere Werte eine bessere Leistung, sodass auch hier das positive Vorzeichen der Erwartung entspricht. Dies gilt ebenso für das positive Vorzeichen für die Kursform, da Studierende, welche einen Kurs auf erhöhtem Anforderungsniveau besucht haben, bessere Testleistungen erzielen. Der Vergleich der standardisierten Koeffizienten zeigt, dass der Einfluss der beiden Schulleistungsmaße gleich stark ist.
Die unstandardisierten Koeffizienten zeigen, dass eine Verbesserung um eine Note im Abitur mit einer Verbesserung um 0,438 Punkten im Testergebnis einhergeht, was einer halben Standardabweichung entspricht. Einen knapp geringeren Einfluss zeigt auch die letzte Mathematiknote; hier geht eine Notenverbesserung mit einem um 0,303 Punkten (3 Punkte bedeuten hier die Verbesserung um eine Note) besseren Testergebnis einher. Der Unterschied zwischen den beiden Kursformen beträgt sogar mehr als eine halbe Standardabweichung.

## 7.8 Varianzanalyse für alle Prädiktoren

Die bisher einzeln betrachteten Faktoren und ihr Einfluss auf die Leistungen im Einstufungstest sollen nun gemeinsam betrachtet werden. Dadurch lassen sich die einzelnen Einflussfaktoren voneinander trennen und ihre jeweiligen Effektstärken vergleichen. Dazu wird eine Varianzanalyse mit den sieben Faktoren Anfangsjahr, Studiengang, Bundesland, Geschlecht, Schuldauer, zeitl. Lücke sowie Kursform, mit den jeweiligen Ausprägungen wie oben, berechnet. Das Ergebnis ist in Tabelle 17 zu sehen.
Dabei ist zu beachten, dass nun aufgrund des Ausschlusses aller Teilnehmenden, für die auch nur ein Wert fehlt, die Gesamtzahl deutlich zurück geht, sodass diese Analyse auf $N = 615$ Studierenden beruht.
Zunächst zeigt sich, dass weiterhin alle Variablen mit Ausnahme des Anfangsjahrs einen signifikanten Beitrag liefern. Dabei besitzt der Studiengang eine große Effektstärke, während Geschlecht und verzögerte Studienaufnahme einen kleinen Effekt besitzen. Die weiteren Effekte sind alle von mittlerer Größe.
Insgesamt erklären die sechs Variablen des Modells $40,8$ % der Varianz innerhalb der Testergebnisse. Eine weitere Aufnahme der Interaktionseffekte liefert keinen wesentliche Verbesserung der Modellanpassung.

| Prädiktor | $df$ | $F$ | $p$ | $\eta^2_{par}$ |
|---|---|---|---|---|
| korrigiertes Modell | 11 | 38,416 | $\leq 0,001$ | |
| konstanter Term | 1 | 256,199 | $\leq 0,001$ | |
| Studiengang | 3 | 34,896 | $\leq 0,001$ | 0,146 |
| Bundesland | 3 | 16,380 | $\leq 0,001$ | 0,074 |
| Schuldauer | 1 | 43,897 | $\leq 0,001$ | 0,067 |
| Verzögerung | 2 | 5,831 | $0,003$ | 0,019 |
| Geschlecht | 1 | 18,672 | $\leq 0,001$ | 0,030 |
| Kursform | 1 | 59,014 | $\leq 0,001$ | 0,088 |

Tabelle 17: Ergebnisse der Varianzanalyse mit 7 Faktoren ($df$: Anzahl der Freiheitsgerade, $F$: F-Wert der Varianzanalyse, $p$: Signifikanzniveau für die Verschiedenheit von 0 des jeweiligen Koeffizienten im Endmodell, $\eta^2_{par}$: partielles $\eta^2$)

Die zugehörigen Post-hoc-Tests der sechs Prädiktorvariablen entsprechen genau den eindimensionalen Vergleichen der vorigen Abschnitte. In den vorherigen Analysen wurden dabei allerdings aufgrund des oben erwähnten fallweisen Ausschlusses größere Gruppen untersucht. Unter den hier untersuchten Studierenden erweisen sich jedoch genau die selben Gruppen als signifikant voneinander verschieden.

## 7.9 Lineare Regression für alle Prädiktoren

In einem letzten Modell werden nun alle Prädiktoren, das bedeutet die Schulleistungsdaten sowie die weiteren Prädiktoren aus Abschnitt 7.8, verwendet, um das Testergebnis vorherzusagen. Dabei werden entsprechend der Beschreibung in Abschnitt 5.5 die ordinalen und nominalen Variablen in Dummy-Variablen überführt. Als Referenzkategorie dienen also Mathematikstudierende für den Studiengang, Niedersachsen für das Bundesland, Männer für Geschlecht, zwölf Jahre für die Schuldauer, direkter Studienbeginn für die Verzögerung sowie grundlegendes Anforderungsniveau für die Kursform. Das Modell ist in Tabelle 18 zu sehen und beruht auf einer Stichprobengröße von $N = 600$ Personen.

| Prädiktor | $B$ | $\beta$ | $t$ | $p$ | Änderung in $R^2$ |
|---|---|---|---|---|---|
| Konstante | 0,930 | | 3,589 | $\leq 0,001$ | |
| Abiturnote | -0,436 | -0,312 | -7,605 | $\leq 0,001$ | 0,375 |
| Kursform | 0,487 | 0,229 | 8,298 | $\leq 0,001$ | 0,068 |
| letzte Mathenote | 0,078 | 0,237 | 5,973 | $\leq 0,001$ | 0,037 |
| Geschlecht_weiblich | -0,330 | -0,164 | -5,866 | $\leq 0,001$ | 0,034 |
| Schuldauer_13Jahre | -0,431 | -0,214 | -7,407 | $\leq 0,001$ | 0,027 |
| Studiengang_LA | -0,350 | -0,149 | -5,061 | $\leq 0,001$ | 0,015 |
| Bundesland_Schl-Hol. | 0,452 | 0,128 | 4,472 | $\leq 0,001$ | 0,013 |
| Bundesland_Hessen | 0,292 | 0,109 | 3,972 | $\leq 0,001$ | 0,011 |
| Studiengang_Info. | -0,199 | -0,089 | -2,788 | 0,005 | 0,005 |

Tabelle 18: Lineare Regressionsanalyse mit dem Testergebnis als abhängiger Variable ($B$: unstandardisierte Regressionskoeffizienten, $\beta$: standardisierte Regressionskoeffizienten, $p$: Signifikanzniveau für die Verschiedenheit von 0 des jeweiligen Koeffizienten im Endmodell, Änderung in $R^2$: Vergrößerung des Bestimmtheitsmaßes durch schrittweises Einfügen des Prädiktors)

Es zeigt sich, dass unter Berücksichtigung der Schulleistungsdaten weiterhin alle anderen Prädiktoren außer der zeitlichen Verzögerung einen signifikanten Beitrag zur linearen Regression leisten. Dabei werden die Dummy-Variablen für das Bundesland Nordrhein-Westfalen und den Studiengang Physik nicht in die Regressionsgleichung aufgenommen. Diese beiden Variablen zeigten bereits in den Varianzanalysen keine Differenz zu den jeweiligen Referenzkategorien. Die beiden Kategorien der zeitlichen Verzögerung hingegen zeigten in diesen Analysen sowie der mehrfaktoriellen Varianzanalyse einen signifikanten Einfluss, der in der linearen Regression nicht bestehen bleibt.
Alle aufgenommenen Variablen zeigen dabei in die gleiche Richtung wie sie es in den einfaktoriellen Analysen getan haben. Insgesamt erklärt das Modell mit $R^2 = 0,585$ knapp 18 % mehr Varianz als das Modell, welches die beiden Zensuren aus der Schule nicht berücksichtigt bzw. 10 % mehr als das Modell, welches ausschließlich die drei Leistungsmaße beinhaltet.
Im Vergleich der beiden Leistungsmaße zeigen die standardisierten Koeffizienten, dass der Einfluss der Abiturnote stärker ist als die letzte Mathenote. Für die Dummy-Variablen geben die unstandardisierten Koeffizienten den Abstand der Gruppe zur Referenzgruppe an, sodass hier die standardisierten Koeffizienten weniger aussagekräftig sind. Hier besitzen die Kursform sowie die Schuldauer die höchsten Beträge. Dennoch liegen alle Koeffizienten in einer ähnlichen Größenordnung.

## 7.10 Teilgebiete

In diesem Abschnitt sollen die acht Testbereiche und ihre Abhängigkeit von den betrachteten Variablen genauer betrachtet werden. Ziel ist es zu untersuchen, ob es gebietsabhängige Unterschiede gibt. Für den Einfluss der Prädiktoren auf die verschiedenen Gebiete werden einzelne Varianzanalysen analog zu dem Vorgehen aus Abschnitt 7.8 durchgeführt. In Tabelle 19 sind die zugehörigen Effektstärken der Prädiktorvariablen zu sehen. Für die Interpretationen werden hier insbesondere die Einordnungen der mathematischen Gebiete auf die Klassenstufen verwendet.

| Variable | Alg | Glei | Poly | Exp | Trig | Vec | Diff | Int |
|---|---|---|---|---|---|---|---|---|
| Studiengang | 0,046 | 0,087 | 0,069 | 0,076 | 0,113 | 0,090 | 0,061 | 0,094 |
| Bundesland | 0,038 | 0,060 | 0,051 | 0,056 | 0,075 | 0,027 | 0,013 | 0,046 |
| Dauer | 0,058 | 0,023 | 0,013 | 0,022 | 0,042 | 0,035 | 0,040 | 0,024 |
| Verzögerung | | | 0,011 | | 0,013 | 0,052 | 0,028 | 0,025 |
| Geschlecht | 0,009 | 0,008 | 0,011 | 0,045 | 0,016 | | 0,012 | 0,008 |
| Kursform | 0,058 | 0,033 | 0,045 | 0,018 | 0,024 | 0,091 | 0,050 | 0,077 |
| $R^2$ | 0,175 | 0,245 | 0,239 | 0,226 | 0,291 | 0,343 | 0,255 | 0,310 |

Tabelle 19: Effektstärken der einzelnen Variablen auf die verschiedenen Bereiche; alle Effekte sind mit $p \leq 0,05$ signifikant

Zunächst zeigt sich, dass mit Ausnahme der verzögerten Studienaufnahme, und für einen Bereich das Geschlecht, alle Variablen in jedem Gebiet einen signifikanten Einfluss auf die Testergebnisse besitzen. Die Varianzaufklärung für die acht Gebiete liegt erwartungsgemäß für alle Gebiete niedriger als für den Gesamttest. Dabei zeigt sich als leichter Trend, dass die Gebiete, welche in der Schulzeit später behandelt werden (siehe dazu auch Abschnitt 5.2), tendenziell höhere Bestimmtheitsmaße besitzen. Diese steigen von 17,5 % im Falle der algebraischen Grundlagen bis zur Vektorrechnung mit 34,3 % an.
Für die Variablen Studiengang, Bundesland, Schuldauer und Geschlecht ergeben sich dabei allerdings zwischen den verschiedenen Gebieten keine systematischen Verschiebungen. Wird die Variable Kursform betrachtet, zeigt sich, dass hier für die letzten drei Gebiete die Effektstärken größer sind als für die fünf Gebiete der Sekundarstufe I. Die zugehörigen Mittelwertunterschiede liegen dennoch über die acht Gebiete hinweg in der gleichen Größenordnung (siehe dazu auch Tabelle 34).
Für die zeitlich verzögerte Studienaufnahme zeigt sich ebenso, dass für die im Unterricht später behandelten Themen die Effektstärken zunehmen bzw. erst einen signifikanten Einfluss besitzen, insbesondere für die Gebiete in der Sekundarstufe II.

| Test | Gruppe | 0 Jahre | 1 Jahr | SuS |
|---|---|---|---|---|
| Grundlagen der Algebra | 1 Jahr | | | |
| | 2 Jahre | $d = 0,30$ | $d = 0,33$ | $d = 0,42$ |
| | SuS | | | |
| Gleichungen und Ungleichungen | 1 Jahr | | | |
| | 2 Jahre | $d = 0,49$ | $d = 0,48$ | |
| | SuS | $d = 0,28$ | | |
| Polynome | 1 Jahr | | | |
| | 2 Jahre | $d = 0,82$ | $d = 0,66$ | $d = 0,75$ |
| | SuS | | | |
| Exponential- und Logarithmus-funktionen | 1 Jahr | | | |
| | 2 Jahre | $d = 0,44$ | $d = 0,37$ | |
| | SuS | $d = 0,34$ | | |
| Trigonometrie | 1 Jahr | $d = 0,23$ | | |
| | 2 Jahre | $d = 0,85$ | $d = 0,61$ | $d = 0,64$ |
| | SuS | | | |
| Vektor-rechnung | 1 Jahr | $d = 0,39$ | | $d = 0,53$ |
| | 2 Jahre | $d = 0,96$ | $d = 0,58$ | $d = 1,14$ |
| | SuS | | | |
| Differential-rechnung | 1 Jahr | | | |
| | 2 Jahre | $d = 0,96$ | $d = 0,82$ | $d = 1,13$ |
| | SuS | | | |
| Integral-rechnung | 1 Jahr | $d = 0,25$ | | |
| | 2 Jahre | $d = 0,90$ | $d = 0,64$ | $d = 0,80$ |
| | SuS | | | |

Tabelle 20: Effektstärken für signifikante Differenzen ($p \leq 0,05$) für die verschiedenen Verzögerungen bzw. Schülerinnen und Schüler; in der Spalte ist die Gruppe mit dem jeweils höheren Mittelwert eingetragen.

Diese Unterschiede sollen nun bezogen auf die drei Gruppen der Studierenden genauer untersucht werden. Auch die Schülerinnen und Schüler werden dabei direkt als vierte Gruppe in die Analyse einbezogen. Hierbei zeigen sich im Gegensatz zur Kursform Unterschiede über die verschiedenen Gebiete hinweg. Dabei ist insbesondere von Interesse, welchen Einfluss zeitliche Verläufe auf die Testergebnisse besitzen.
Für alle acht Bereiche ergibt sich in der einfaktoriellen Varianzanalyse ein signifikanter Einfluss der Verzögerung auf das Ergebnis in der Testleistung. Die Effektstärken der signifikanten Post-hoc-Gruppenvergleiche sind in Tabelle 20 zu sehen. Die Tabelle 20 zeigt zunächst, dass in allen Bereichen die Unterschiede zwischen den Studierenden ohne verzögerten Studienbeginn und den Studierenden mit mindestens zwei Jahren Verzögerungen signifikant sind. Dabei ergeben sich für die Bereiche Grundlagen, Gleichungen und Ungleichungen sowie Exponential- und Logarithmusrechnung kleine und für die

weiteren fünf Bereiche starke Effekte.
Ebenso sind alle Unterschiede zwischen den Studierenden mit genau einem Jahr Verzögerung und den Studierenden mit mindestens zwei Jahren Verzögerung signifikant. Hierbei liegen die Effektstärken in den gleichen drei Bereichen Grundlagen, Gleichungen und Ungleichungen sowie Exponential- und Logarithmusrechnung in einem kleinen Bereich, und die anderen fünf Bereiche zeigen mittlere Effekte.
Im Vergleich der Studierenden ohne verzögerten Studienbeginn und mit einem Jahr Verzögerung ergeben sich signifikante Unterschiede in den drei Bereichen Trigonometrie, Vektorrechnung und Integralrechnung jeweils mit kleiner Effektstärke.
Im Vergleich der Schülerinnen und Schüler ergeben sich für sechs der Bereiche signifikante Unterschiede zu den Studierenden mit mindestens zwei Jahren Verzögerungen. Dabei zeigt der Bereich Grundlagen einen kleinen Effekt, die Bereiche Polynome, Trigonometrie und Integralrechnung jeweils einen mittleren Effekt sowie die Bereiche Vektorrechnung und Differentialrechnung einen starken Effekt.
Im Vergleich der Schülerinnen und Schüler gibt es drei kleine signifikante Effekte in den Bereichen Exponential- und Logarithmusrechnung sowie Vektorrechnung und Integralrechnung, in denen die Schülerinnen und Schüler jeweils geringere Leistungen zeigen. Nur in einem Bereich zeigen die Schülerinnen und Schüler im Vergleich zu den Studierenden mit genau einem Jahr Verzögerung eine signifikant höhere Leistung; dies ist ein kleiner Effekt im Bereich Vektorrechnung.
Insgesamt betrachtet zeigt sich, dass über alle acht Bereiche hinweg im Vergleich mit den Studierenden mit mindestens zwei Jahren Verzögerungen die anderen drei Gruppen durchgängig signifikant höhere Leistungen erzielen. Dabei unterscheidet sich dennoch die Stärke zwischen den Gruppen: Während die Studierenden ohne Verzögerung den größten Abstand besitzen, sind die Effekte für die Schülerinnen und Schüler sowie die Gruppe mit einem Jahr Verzögerung von jeweils ähnlicher Höhe.
Gleichzeitig zeigt sich, dass die Effektstärken in den Bereichen, welche in der Sekundarstufe II behandelt werden, also Differentialrechnung, Integralrechnung sowie Vektorrechnung, höher sind als in den fünf früher behandelten Bereichen. Zu den späteren Bereichen gehören auch die drei Bereiche, Trigonometrie, Vektorrechnung sowie Integralrechnung, welche einen Unterschied zwischen Studierenden mit einem Jahr Verzögerung und unverzögert beginnenden Studierenden besitzen, während die Unterschiede zu den Schülerinnen und Schülern in den früher behandelten Bereichen Gleichungen und Ungleichungen sowie Exponential- und Logarithmusfunktionen liegen.

# 8 Längsschnittliche Auswertung der Klausurergebnisse

In diesem Abschnitt sollen die Klausurergebnisse am Ende des ersten Semesters untersucht werden. Dies geschieht, wie in Abschnitt 5.5 beschrieben, für die Teilnehmenden des Propädeutikums aus den Jahrgängen 2013 und 2014.
Dabei werden aus der Stichprobe jeweils nur die Studiengänge untersucht, für die diese Veranstaltungen verpflichtend waren. Dies bedeutet für das Jahr 2013 die Bachelorstudierenden in Mathematik und Physik sowie die Lehramtsstudierenden. Für das Jahr 2014 sind diese Veranstaltungen für die Bachelorstudierenden in Physik nicht Pflicht, sodass auch diese nicht in die Untersuchung aufgenommen werden.
Im ersten Abschnitt wird zunächst dargestellt, welcher Anteil an Studierenden aus der Stichprobe an den Klausuren am Ende des ersten Semesters teilgenommen hat, und, wie sich diese von den Nichtteilnehmenden im Hinblick auf ihre Testleistung im Propädeutikum unterscheiden. Im zweiten Schritt wird dann die Frage des Bestehens oder Durchfallens der Klausuren dichotom untersucht und anschließend auch die Untersuchung der Klausurzensuren.
Diese Untersuchungen beruhen auf logistischen bzw. linearen Regressionen unter Verwendung der Prädiktoren, welche im Propädeutikum erhoben wurden. Insbesondere das Verhältnis der drei Leistungsmaße Abiturnote, letzte Mathenote und Testergebnis zu Beginn des Studiums zu den Klausurergebnisse ist dabei von Interesse. Es werden also ausschließlich Variablen der Studierenden verwendet, welche vor dem Studienbeginn erhoben worden sind, um die Ergebnisse nach dem ersten Semester vorherzusagen.

## 8.1 Deskriptive Beschreibung der Klausurergebnisse

Um eine Übersicht über die untersuchten Klausuren zu erhalten, werden zunächst die Klausurtermine einzeln dargestellt. Dabei wird insbesondere betrachtet, ob es Unterschiede in Bestehensquoten und Zensuren zwischen Teilnehmenden des Propädeutikums und den Studierenden gibt, die nicht am Propädeutikum teilgenommen haben. Es wird also die beschriebene Stichprobe mit den weiteren Studienanfängerinnen und -anfängern verglichen.
In Tabelle 21 sind dazu die deskriptiven Daten der untersuchten Klausuren nach den Klausurterminen für alle Klausurteilnehmenden zu sehen. Die Durchschnittszensuren beziehen sich dabei nur auf die Studierenden, welche die jeweilige Klausur bestanden haben.
Es zeigen sich Bestehensquoten, die zwischen 40 % und 68 % liegen. Dabei zeigen sich für die Klausuren in "Differential und Integralrechnung" größere Schwankungen zwischen den beiden Klausurterminen. Für die durchschnittliche Klausurzensur, wenn bestanden wurde, ergibt sich mit einer Ausnahme ein Durchschnitt, der bei ungefähr 3,0 oder etwas darüber liegt.

| Klausurtermin | Teilnehmende | Bestehens-quote | Durchschnittszensur (wenn bestanden) |
|---|---|---|---|
| Diff1-WS13/14 | 142 | 96 (68 %) | 3,12 |
| Diff1-SS14 | 80 | 37 (46 %) | 3,57 |
| Diff1-WS14/15 | 103 | 63 (61 %) | 2,06 |
| Diff1-SS15 | 42 | 17 (40 %) | 2,92 |
| Agla1-WS13/14 | 199 | 106 (53 %) | 2,85 |
| Agla1-SS14 | 123 | 71 (58 %) | 3,48 |
| Agla1-WS14/15 | 109 | 67 (61 %) | 2,98 |
| Agla1-SS15 | 66 | 34 (52 %) | 3,69 |

Tabelle 21: Anzahl der Teilnehmenden, Bestehensquote und Durchschnittszensuren der bestandenen Klausuren einzeln für die Klausuren des ersten Semesters, bezogen auf alle Teilnehmenden

Für den Vergleich zwischen allen Studierenden, welche die Klausuren geschrieben haben, und den Teilnehmenden aus dem Propädeutikum zeigt Tabelle 22 die entsprechenden Daten bezogen ausschließlich auf die Teilnehmenden des Propädeutikums.

| Klausurtermin | Teilnehmende | Bestehens-quote | Durchschnittszensur (wenn bestanden) |
|---|---|---|---|
| Diff1-WS13/14 | 71 | 45 (63 %) | 3,09 |
| Diff1-SS14 | 43 | 21 (49 %) | 3,53 |
| Diff1-WS14/15 | 44 | 29 (66 %) | 2,09 |
| Diff1-SS15 | 8 | 3 (38 %) | 2,47 |
| Agla1-WS13/14 | 90 | 44 (49 %) | 2,89 |
| Agla1-SS14 | 56 | 30 (54 %) | 3,39 |
| Agla1-WS14/15 | 40 | 27 (68 %) | 3,02 |
| Agla1-SS15 | 18 | 7 (39 %) | 3,61 |

Tabelle 22: Anzahl der Teilnehmenden, Bestehensanteil und Durchschnittszensuren der bestandenen Klausuren einzeln für die Klausuren des ersten Semesters, bezogen auf die Teilnehmenden des Propädeutikums

Im Vergleich der beiden Tabellen ergeben sich nahezu keine Unterschiede zwischen den beiden Gruppen. Dies gilt sowohl für die Bestehensquote als auch für die Durchschnittsnote der bestandenen Klausuren. Die Teilnehmenden der Stichprobe an den Klausuren erzielen also sehr ähnlich Klausurresultate wie die Klausurteilnehmenden, die nicht am Propädeutikum teilgenommen haben.
Im Vergleich der beiden Veranstaltungen liegen die Bestehensquoten der vier Prüfungen in "Analytische Geometrie und lineare Algebra 1" enger zusammen als für die Prüfungen in "Differential und Integralrechnung". Ebenso fallen Unterschiede in den Durchschnittszensuren höher aus.

## 8.2 Beschreibung der Klausurteilnehmenden

In diesem Abschnitt werden nun wieder ausschließlich die Teilnehmenden am Propädeutikum, wie in den vorigen Abschnitten, betrachtet. Dabei wird untersucht, wie groß der Anteil von ihnen ist, der auch an einer Klausur am Ende des ersten Semesters teilnimmt. Dafür werden hier nur die Teilnehmenden aus dem Propädeutikum einbezogen, für welche die beiden untersuchten Veranstaltungen verpflichtend waren. Am Ende wird untersucht, ob es Unterschiede in den Leistungsmaßen zwischen den Teilnehmenden gibt, welche an einer Klausur teilgenommen haben, und denen, welche an keiner der Klausuren teilgenommen haben.

Für die Klausurteilnahme am Ende des ersten Semesters ergeben sich getrennt nach den Jahrgängen die Daten aus Tabelle 23.

| | Studierendengruppe | 2013 | 2014 |
|---|---|---|---|
| | Propädeutikumsteilnehmende | 200 | 214 |
| davon | aus den betrachteten Studiengängen | 159 | 61 |
| davon | mindestens eine Klausur geschrieben | 106 | 43 |
| | Anteil | 67 % | 70 % |

Tabelle 23: Anteil der Teilnehmenden an mindestens einer Klausur am Ende des ersten Semesters

Es haben also in beiden Jahrgängen knapp 70 % der Teilnehmenden aus dem Propädeutikum an mindestens einer der Klausuren teilgenommen. Die auffällige Verschiebung in den absoluten Zahlen ergibt sich aus der beschriebenen Veränderung des Physik-Studiengangs. Es zeigt sich allerdings keine Verschiebung zwischen den beiden Jahrgängen, welche mit einer veränderten Teilnahme an Studiengängen einherging, in den Quoten.

Nun werden unter Verwendung eines $t$-Test die späteren Klausurteilnehmenden mit den Studierenden, die an keiner der Klausuren teilgenommen haben, im Bezug auf ihre Testleistungen vor dem Studium verglichen. Dabei ergibt sich für das Jahr 2013 mit $t(157) = 3,59$; $p \leq 0,001$ und $d = 0,61$ eine signifikante Differenz zwischen den beiden Gruppen. Es besitzen die Studierenden, welche eine Klausur geschrieben haben, mit $\mu = 1,21$ ($\sigma = 0,75$) einen höheren Mittelwert im Testergebnis als die Studierenden, welche an keiner Klausur teilgenommen haben, die einen Mittelwert von $\mu = 0,75$ ($\sigma = 0,77$) besitzen.

Ein ebenso mittlerer Effekt ergibt sich für das Jahr 2014 mit $t(59) = 2,05$; $p = 0,045$ und $d = 0,57$, für den die Mittelwerte $\mu = 1,29$ ($\sigma = 0,90$) und $\mu = 0,79$ ($\sigma = 0,84$) ergeben. Das Testergebnis zeigt hier also einen wesentlichen Einfluss auf das Antreten zu mindestens einer Klausur am Ende des ersten Semesters.

## 8.3 Bestehen der Klausuren

Für die Betrachtung des Bestehens der Klausuren im ersten Semester werden nun logistische Regressionen durchgeführt. Dafür werden alle erhobenen Variablen, wie in Abschnitt 7, sowie die Leistung im Einstufungstest als Prädiktoren verwendet. Dabei zeigen diese Analysen unter Verwendung aller Prädiktoren keine signifikanten Einflüsse von den nicht leistungsbezogenen Variablen (siehe dazu auch Tabelle 35 im Anhang). Es bleiben also nur die drei Variablen "Testergebnis", "Abiturnote" und "letzte Mathenote".
Da durch teilweise unvollständige Angaben zu den weiteren Prädiktoren die Stichprobengrößen verkleinert werden, werden im Folgenden Ergebnisse dargestellt, deren Analysen sich auf alle Personen beziehen, für welche diese drei Prädiktoren erhoben wurden, um die Anzahl der Fälle möglichst groß zu halten. Weiterhin wird das Anfangsjahr als vierte Prädiktorvariable in die Analyse mit aufgenommen, da in diesem Fall davon auch die besuchte Veranstaltung sowie die ersten möglichen Klausurtermine abhängen. In dieser Analyse werden dabei nur Ergebnisse von Klausuren betrachtet, welche im ersten Semester geschrieben wurden.
Für die Betrachtung der Veranstaltung "Differential- und Integralrechnung 1" ergeben sich unter schrittweisem Hinzufügen der Variablen die Modelle 1 und 2 in Tabelle 24. Die weiteren Variablen werden nicht aufgenommen, da sie keine signifikante Verbesserung des Modells liefern. Als Vergleich ergeben sich die Modelle 3 und 4 unter Ausschluss des Testergebnisses als Prädiktor, sodass nur Schulleistungen verwendet werden.
Es zeigt sich, dass das Modell 2, welches das Testergebnis sowie den Jahrgang als Prädiktoren beinhaltet, die beste Modellpassung besitzt. Dieses Modell besitzt mit einem Mc Faddens $R^2$ von 0,27 ein als gut zu betrachtendes Pseudo-Bestimmtheitsmaß, der Informationsindex AIC besitzt den geringsten Wert und es ordnet $85,1\%$ der Studierenden ein korrektes Testergebnis zu. Insbesondere zeigt sich, dass das Testergebnis als Leistungsmaß eine höhere Prädiktionskraft besitzt als die Leistungsmaße der Schule, für welche die letzte Mathenote den besten Prädiktor darstellt. Das entsprechende Modell 4 besitzt eine korrekte Zuordnung von $79,1\%$ der Prüfungsteilnehmenden.

| | Modell 1 | Modell 2 | Modell 3 | Modell 4 |
|---|---|---|---|---|
| Konstante | -0,436 | 0,035 | -4,303 | -4,546 |
| Odds-ratio | 0,65 | 1,04 | 0,01 | 0,01 |
| Testergebnis | 1,589 | 1,880 | | |
| Odds-ratio | 4,90 | 6,55 | | |
| Abiturnote | | | | |
| Odds-ratio | | | | |
| letzte Mathenote | | | 0,445 | 0,514 |
| Odds-ratio | | | 1,56 | 1,67 |
| Jahrgang | | -1,799 | | -1,446 |
| Odds-ratio | | 0,17 | | 0,248 |
| N | 134 | 134 | 134 | 134 |
| korrekt zugeordnete Fälle | 79,9 % | 85,1 % | 77,6 % | 79,1 % |
| AIC | 116,63 | 106,27 | 127,58 | 120,35 |
| Mc Fadden $R^2$ | 0,18 | 0,27 | 0,10 | 0,17 |

Tabelle 24: Logistische Regressionen zur Bestimmung des Erfolgs in der Klausur "Differential- und Integralrechnung 1" in Abhängigkeit von Abiturdurchschnitt, letzter Mathenote, Testergebnis sowie Anfangsjahr; alle Modelle zeigen eine signifikant bessere Passung als das konstante Modell mit $p \leq 0,001$; der Trennwert der Zuordnung beträgt 0,5

Entsprechend dem Vorgehen für die Klausur in "Differential- und Integralrechnung 1" werden unter Verwendung alle Prädiktoren mit der Einschluss-Methode nacheinander die Modelle 1 und 2 für die Klausur in "Analytische Geometrie und Lineare Algebra 1" berechnet. Wird wiederum die hierbei zunächst verwendete Abiturnote aus der Analyse ausgeschlossen, so ergibt sich schrittweise das Modell 3, in dem alle drei weiteren Prädiktoren einen signifikanten Beitrag leisten. Als Modell, in dem die Verwendung des Testergebnisses ausgeschlossen wird, ergibt sich wiederum das Modell 1.
Im Vergleich der drei Modelle zeigt das Modell 3 in allen drei Maßen mit 78,9% korrekt zugeordneten Fällen, dem geringsten Informationsindex und einem guten Mc Faddens $R^2$ die beste Modellpassung. Dabei besitzt dieses Modell wiederum als Leistungsmaß das Testergebnis, in diesem Fall aber zusätzlich auch die letzte Mathenote. Dennoch zeigt das Modell 1, dass das Abiturergebnis hier den stärksten Einzelprädiktor darstellt und das Modell 2 nur eine knapp schwächere Passung liefert als das Modell 3.

| | Modell 1 | Modell 2 | Modell 3 |
|---|---|---|---|
| Konstante | 4,414 | 2,016 | -3,404 |
| Odds-Ratio | 82,60 | 7,51 | 0,3 |
| Testergebnis | | 0,924 | 1,264 |
| Odds-Ratio | | 2,52 | 3,54 |
| Abiturnote | -1,892 | -1,161 | |
| Odds-Ratio | 0,15 | 0,31 | |
| letzte Mathenote | | | 0,260 |
| Odds-Ratio | | | 1,30 |
| Jahrgang | | | -1,007 |
| Odds-Ratio | | | 0,37 |
| N | 141 | 141 | 141 |
| korrekt zugeordnete Fälle | 75,2% | 78,7% | 78,9% |
| AIC | 146,56 | 141,15 | 140,68 |
| Mc Fadden $R^2$ | 0,14 | 0,18 | 0,20 |

Tabelle 25: Logistische Regressionen zur Bestimmung des Erfolgs in der Klausur "Analytische Geometrie und lineare Algebra 1" in Abhängigkeit von Abiturdurchschnitt, letzter Mathenote, Testergebnis sowie Anfangsjahr; alle Modelle zeigen signifikant bessere Passung als das konstante Modell mit $p \leq 0,001$

Für beide Vorlesungen zeigt sich, dass das Testergebnis als Prädiktor für das jeweils beste Modell verwendet wird, im Fall der "Analytische Geometrie und lineare Algebra 1" leistet die letzte Mathenote einen zusätzlichen Beitrag. Ebenso wird in beiden Modellen der Jahrgang aufgenommen, dabei ist zu bedenken, dass in einem der Jahre die Physikstudierenden Teil der Stichprobe waren, während dies für den zweiten Jahrgang nicht zutrifft.
Um nun den Einfluss der veränderten Teilnahme an den Klausuren zu untersuchen, werden die Analysen wiederholt, wobei die Physikstudierenden aus den Analysen ausgeschlossen werden. Dadurch ergeben sich in beiden Jahrgängen die gleichen Studierendengruppen, nämlich ausschließlich die Studiengänge Mathematik und die Lehramtsstudierenden. Es ergeben sich unter schrittweisem Einschluss die Modelle aus Tabelle 26.

| | Diff 1 | Agla 1 |
|---|---|---|
| Konstante | 0,906 | 5,646 |
| Odds-ratio | 2,47 | 283,16 |
| Testergebnis | 2,160 | |
| Odds-ratio | 8,67 | |
| Abiturnote | | -2,428 |
| Odds-ratio | | 0,09 |
| letzte Mathenote | | |
| Odds-ratio | | |
| Jahrgang | -3,112 | |
| Odds-ratio | 0,04 | |
| N | 66 | 65 |
| korrekt zugeordnete Fälle | 84,8% | 72,3% |
| AIC | 56,13 | 65,129 |
| Mc Fadden $R^2$ | 0,35 | 0,24 |

Tabelle 26: Logistische Regressionen zur Bestimmung des Erfolgs in beiden Klausuren in Abhängigkeit von Abiturdurchschnitt, letzter Mathenote, Testergebnis sowie Anfangsjahr ohne die Physikstudierenden; alle Modelle zeigen signifikant bessere Passung als das konstante Modell mit $p \leq 0,001$

Es zeigt sich, dass für die Klausur zur "Analytische Geometrie und lineare Algebra 1" der Effekt des Jahrgangs verschwindet, während er für "Differential- und Integralrechnung 1" erhalten bleibt und sogar zunimmt. Gleichzeitig steigt das Pseudo-Bestimmtheitsmaß im Vergleich zu der Betrachtung aller Studierenden wesentlich an, während die korrekte Zuordnung etwas abnimmt. Somit ergibt sich insgesamt ein uneinheitliches Bild für die Abhängigkeit vom Jahrgang für die beiden Vorlesungen, insbesondere lassen sich Unterschiede nicht alleine auf die Veränderung in der Zusammensetzung der Klausurschreibenden zurückführen.

## 8.4 Zensuren der Klausuren

Nach der Frage des Bestehens der Klausuren wird nun untersucht, welcher Zusammenhang zwischen den Zensuren und den erhobenen Prädiktoren besteht. Dabei werden in diesem Abschnitt für die Klausurergebnisse jeweils die Zensuren in der ersten geschriebenen Klausur und die letzte Klausurnote unterschieden.
Dazu werden zunächst die Korrelationen zwischen den, auch oben verwendeten, Leistungsmaßen zu den Klausurergebnissen berechnet. Anschließend werden lineare Regressionen für die Klausurergebnisse unter Verwendung aller Prädiktoren, die erhoben wurden, durchgeführt.

Bei der Berechnung der Korrelationen zeigt sich, dass diese für alle Kombinationen in einem mittleren bis hohen Bereich liegen. Die genauen Werte sind in Tabelle 27 zu sehen. Es ergibt sich dabei für alle Klausurergebnisse ein einheitliches Bild bezüglich der Reihenfolge der Korrelationen zu den drei Leistungsmaßen. In allen Fällen zeigt das Testergebnis die höchste Korrelation, gefolgt von der Abiturnote und der letzten Mathenote.

| | Testergebnis | Abiturnote | letzte Mathenote |
|---|---|---|---|
| Diff_ersteNote | -0,629 | 0,466 | -0,444 |
| | N=146 | N=142 | N=135 |
| Diff_letzte Note | -0,578 | 0,404 | -0,379 |
| | N=147 | N=143 | N=136 |
| Agla_ersteNote | -0,588 | 0,517 | -0,478 |
| | N=150 | N=147 | N=139 |
| Aga_letzte Note | -0,547 | 0,517 | -0,442 |
| | N=153 | N=150 | N=142 |

Tabelle 27: Korrelationen zwischen den Klausurergebnissen und den Leistungsmaßen vor dem Studienbeginn; alle Korrelationen sind signifikant mit $p \leq 0,001$; Abweichungen der Fallzahlen ergeben sich aus der Definition der Variablen, wie in Abschnitt 5 beschrieben.

Im Weiteren werden nun die Klausurergebnisse mit Hilfe linearer Regressionen untersucht. Dabei zeigt sich wiederum, dass die nicht leistungsbezogenen Variablen keinen signifikanten Einfluss auf die Klausurergebnisse besitzen (siehe dazu auch Tabelle 36 und Tabelle 37). Daher werden wiederum die Regressionen ausschließlich mit den vier Variablen durchgeführt, die auch für die logistischen Regressionen verwendet wurden.

| Prädiktor | $B$ | $\beta$ | $t$ | $p$ | Änderung in $R^2$ |
|---|---|---|---|---|---|
| erste Klausur: | | | | | |
| Konstante | 4,016 | | 7,862 | $\leq 0,001$ | |
| Testergebnis | -0,815 | -0,483 | -5,786 | $\leq 0,001$ | 0,414 |
| Anfangsjahr | -0,709 | -0,256 | -3,785 | $\leq 0,001$ | 0,040 |
| Abiturnote | 0,552 | 0,210 | 2,502 | $0,014$ | 0,025 |
| letzte Klausur: | | | | | |
| Konstante | 4,485 | | 26,264 | $\leq 0,001$ | |
| Testergebnis | -0,876 | -0,573 | -8,034 | $\leq 0,001$ | 0,328 |

Tabelle 28: Lineare Regression für die Klausuren zur "Differential- und Integralrechnung 1"

Für die Veranstaltung "Differential- und Integralrechnung1" ergibt sich für die erste Klausur ein Modell, welches als Prädiktorvariablen mit signifikanten Beiträgen das Testergebnis, das Anfangsjahr sowie die Abiturnote enthält.

Für die letzte Klausur zeigt nur das Testergebnis einen signifikanten Einfluss. Die beiden Modelle basieren jeweils auf einer Stichprobe von $N = 134$ Fällen und sind in Tabelle 28 dargestellt.
Für die erste Klausur ergibt sich mit $R^2 = 0,479$ ein Bestimmtheitsmaß, welches wesentlich höher liegt als das Bestimmtheitsmaß von $R^2 = 0,328$ für die letzte Klausur. Dies entspricht einer um 15 % höheren Varianzaufklärung in der Klausurnote für die erste geschriebene Klausur. Zu einem Teil ist dieser Unterschied durch die signifikanten Beiträge der Variablen Anfangsjahr sowie Abiturnote zu erklären, wobei bereits das Testergebnis alleine einen höheren Beitrag für die erste Klausur liefert als für die letzte Klausur. Das Einschließen des Anfangsjahrs für die erste Klausur entspricht dabei dem Ergebnis der logistischen Regression, insbesondere auch in seiner Richtung.
Der unstandardisierte Regressionskoeffizient für das Testergebnis liegt für beide Modelle in ähnlicher Größenordnung und zeigt, dass eine Verbesserung um einen Punkt, bei einer Standardabweichung von $0,93$, zu einer Verbesserung um eine Note führt. Für die Abiturnote ergibt sich eine Verbesserung, die etwas mehr als halb so hoch ausfällt, in der Klausur. Ebenso für die erste Klausur zeigt sich ein Unterschied von $0,7$ für die beiden Jahrgänge. Der Vergleich der standardisierten Regressionskoeffizienten zeigt, dass in beiden Fällen der Einfluss des Testergebnisses etwa doppelt so stark ist wie der Einfluss der beiden anderen Variablen.

| Prädiktor | $B$ | $\beta$ | $t$ | $p$ | Änderung in $R^2$ |
|---|---|---|---|---|---|
| erste Klausur: | | | | | |
| Konstante | 5,676 | | 6,430 | $\leq 0,001$ | |
| Testergebnis | -0,564 | -0,391 | -4,577 | $\leq 0,001$ | 0,352 |
| letzte Mathenote | -0,128 | -0,197 | -2,468 | $0,015$ | 0,049 |
| Abiturnote | 0,398 | 0,177 | 1,987 | $0,049$ | 0,017 |
| letzte Klausur: | | | | | |
| Konstante | 6,145 | | 11,243 | $\leq 0,001$ | |
| Testergebnis | -0,630 | -0,476 | -6,229 | $\leq 0,001$ | 0,305 |
| Abiturnote eingefügt | | | | | 0,043 |
| letzte Mathenote | -0,142 | -0,237 | -3,134 | $0,002$ | 0,020 |
| Anfangsjahr | 0,445 | 0,203 | 2,994 | $0,003$ | 0,025 |
| Abiturnote entfernt | | | | | -0,006 |

Tabelle 29: Lineare Regression für die Klasuren zur "Analytische Geometrie und lineare Algebra 1"

Für die Regression der Klausuren in "Analytische Geometrie und lineare Algebra 1" beinhaltet das Modell für die erste Klausur alle drei Leistungsmaße, das Anfangsjahr wird hingegen nicht in das Modell aufgenommen. Für die letzte Klausur werden bei schrittweisem Vorgehen zunächst alle Variablen

in das Modell aufgenommen, jedoch die Abiturnote am Ende wieder ausgeschlossen, sodass sich insgesamt auch hier drei Prädiktoren als signifikant erweisen. Die endgültigen Modelle sind in Tabelle 29 zu sehen.
Es zeigt sich auch in diesem Fall, dass mit $R^2 = 0,418$ das Bestimmtheitsmaß für die erste Klausur über dem der letzten Klausur von $R^2 = 0,388$ liegt, obwohl in diesem Fall der Unterschied nur gering ausfällt. Im Unterschied zur Klausur in "Differential- und Integralrechnung 1" liefert in diesen Modellen die letzte Mathenote jeweils einen signifikanten Beitrag. Der Beitrag des Jahrgangs fällt etwas kleiner aus und besitzt ein umgekehrtes Vorzeichen, außerdem tritt er in den Modellen zur "Analytische Geometrie und lineare Algebra 1" in der letzten Klausur und nicht für die erste Klausur auf.
Für diese beiden Modelle ergibt sich mit unstandardisierten Regressionskoeffizienten von $0,6$ eine geringere Verbesserung in der Klausurnote bei Verbesserung im Testergebnis im Vergleich zu "Differential- und Integralrechnung 1". Dazu kommen im Fall der ersten Klausur noch der Einfluss von Abiturnoten sowie letzter Mathenote hinzu, welche zu Veränderungen in der gleichen Größe führen. Auch hier zeigt der Vergleich der standardisierten Regressionskoeffizienten einen ungefähr doppelt so starken Einfluss des Testergebnisses im Vergleich mit den beiden anderen Variablen.
Zur Untersuchung des Effekts, den der Anfangsjahrgang aufgrund verschiedener Veranstaltungen auf die Leistungen spielt, können für die beiden Jahrgänge unabhängige Modelle berechnet werden. Diese einzelnen Modelle liefern dabei allerdings keine Verbesserung gegenüber dem gemeinsamen Modell. Dies gilt ebenso für den Einbezug des Studiengangs als weitere Variable oder das Einschränken auf die Studierenden der Mathematik und des Mathematik Lehramts ohne die Physikstudierenden. Ebenso ergeben sich keine Verbesserungen der Modelle, wenn, wie in (Sorge et al., 2016), nur Studierende betrachtet werden, welche die Klausuren bestanden haben. Es zeigt sich also auch hier, dass eine gemeinsame Modellierung der beiden erhobenen Jahrgänge möglich und vorteilhaft ist.
Für die Klausuren stellt also das Testergebnis den besten Prädiktor für die Klausurnote dar. Dabei liefern die beiden Schulzensuren jeweils kleine zusätzliche Beiträge für die Erklärung der Klausurzensuren. Dies entspricht den Ergebnissen der vorgestellten vorigen Untersuchungen. Für die Zensuren zeigt sich hier jedoch kein so starker Einfluss des Anfangsjahres wie es für das Bestehen der Fall war.

## 8.5 Testgebiete

In diesem Abschnitt soll nun untersucht werden, welche Ergebnisse eine Betrachtung der einzelnen Testgebiete im Zusammenhang mit den Klausurergebnissen liefert. Dazu werden zunächst die Korrelationen der Gebiete mit den Klausuren berechnet.
Die gefundenen Korrelationen nehmen insgesamt Werte an, welche im We-

sentlichen im Bereich zwischen $0,32$ und $0,48$ liegen. Dabei liegen alle Korrelationen, unabhängig vom Gebiet und der Klausur, unter der jeweiligen Korrelation zwischen Gesamtergebniss und der Klausur (siehe dazu auch Tabelle 27). Durch die Schwankungen ergibt sich für die Klausuren in "Analytische Geometrie und lineare Algebra 1", dass die Abiturnote mit $0,52$ eine höhere Korrelation zum Klausurergebniss besitzt als jeder der einzelnen Testbereiche. Im Gegensatz dazu gibt es für "Differential- und Integralrechnung 1" immer noch einzelne Testbereiche, welche eine höhere Korrelationen zeigen, als die Abiturnote. Die letzte Mathematiknote liegt mit ihrer Korrelation jeweils im mittleren Bereich der hier berechneten Korrelationen.

| Testgebiet | Diff ersteNote | Diff letzte Note | Agla ersteNote | Aga letzte Note |
|---|---|---|---|---|
| Alg | -0,427<br>N=146 | -0,327<br>N=147 | -0,265<br>N=150 | n.s. |
| Glei | -0,357<br>N=146 | -0,341<br>N=147 | -0,475<br>N=150 | -0,405<br>N=153 |
| Poly | -0,416<br>N=146 | -0,390<br>N=147 | -0,460<br>N=150 | -0,496<br>N=153 |
| Exp | -0,491<br>N=146 | -0,466<br>N=147 | -0,471<br>N=150 | -0,465<br>N=153 |
| Trig | -0,420<br>N=146 | -0,322<br>N=147 | -0,425<br>N=150 | -0,362<br>N=153 |
| Vec | -0,487<br>N=146 | -0,447<br>N=147 | -0,380<br>N=150 | -0,352<br>N=153 |
| Diff | -0,535<br>N=146 | -0,554<br>N=147 | -0,486<br>N=150 | -0,494<br>N=153 |
| Int | -0,241<br>N=146 | -0,258<br>N=147 | -0,310<br>N=150 | -0,331<br>N=153 |

Tabelle 30: Korrelationen zwischen den Klausurergebnissen und Ergebnissen in den einzelnen Testbereichen; alle Korrelationen sind signifikant mit $p \leq 0,003$; Abweichungen der Fallzahlen ergeben sich aus der Definition der Variablen, wie in Abschnitt 5 beschrieben.

Darüber hinaus zeigen sich keine systematischen Unterschiede zwischen den beiden Klausuren. Werden Schulgebiete betrachtet, welche inhaltlich passend zu den Universitätsklausuren sind, ergeben sich insbesondere auch keine höheren Korrelationen.
Insgesamt zeigt sich also, dass die Betrachtung des Gesamtergebnisses die höchsten Zusammenhänge mit den Klausurergebnissen liefert. Eine Betrachtung der einzelnen Testgebiete zeigt hier, im Gegensatz zu den Testergebnissen vor dem Studium, keine systematischen Auffälligkeiten. Entsprechend werden hier auch keine Ergebnisse für Regressionen mit Hilfe der Ergebnisse

einzelner Gebiete dargestellt, da diese für Vorhersagen schlechter geeignet sind Dies gilt sowohl für lineare als auch logistische Regressionen (siehe dazu auch Tabelle 38 bis Tabelle 41 ).

# 9 Diskussion

Die Ergebnisse werden nun, eingeteilt in drei Abschnitte, diskutiert. Zunächst werden die Ergebnisse des Einstufungstests betrachtet in ihrer Abhängigkeit der verschiedenen Variablen, erst einzeln und dann gemeinsam. Danach werden die Erfolge in den Klausuren des ersten Semesters betrachtet. Am Ende steht die Diskussion der Ergebnisse zu den verschiedenen Testbereichen; diese beziehen sich dabei auf beide Aspekte, die vorher beschrieben wurden.

## 9.1 Testergebnisse

Im ersten Abschnitt werden die Variablen untersucht, welche das Testergebnis beeinflussen. Dabei folgt die Diskussion den einzelnen Forschungsfragen aus Abschnitt 4.

### 9.1.1 Studiengang - (F2)

Die Ergebnisse zu den Differenzen zwischen den betrachteten Studiengängen zeigten in Abschnitt 7.2 zwei Paare. Dabei erzielten die Studierenden in Mathematik und in Physik wesentlich bessere Testleistungen als die Studierenden in angewandter Informatik und die Lehramtsstudierenden. Insbesondere die Deutlichkeit der beiden Differenzen entspricht dabei nicht den vorher vorgestellten Ergebnissen aus Abschnitt 3, da sich dort keine bis geringe Unterschiede in den Leistungen zeigten. Im Vergleich der Mathematik- und der Physikstudierenden war ein so gleiches Ergebnis hingegen erwartbar. Mögliche Erklärungen der Unterschiede zu den anderen beiden Studiengängen könnten sich dabei unterscheiden.
Im Vergleich der Mathematikstudierenden und den Lehramtsstudierenden haben die beschriebenen Studien, welche sich fachspezifische Unterschiede angeschaut haben, insgesamt die mathematisch-naturwissenschaftlichen Lehramtsstudierenden mit den zugehörigen Nicht-Lehramt-Studierenden verglichen; einzelne Fächer wurden nicht betrachtet. Die Unterscheidung der Studierenden nach ihren Fachgebieten wurde dabei gerade damit begründet, dass fachspezifische Unterschiede existieren, welche den globalen Nachteil der Lehramtsstudierenden begründen. Es ist daher nicht unplausibel, wenn es auch zwischen den mathematisch-naturwissenschaftlichen Lehramtsstudierenden weitere Differenzeffekt gibt. Dabei sind auch verschiedene institutionelle Rahmenbedingungen im Vergleich der Fächer zu beachten. So besitzt der Studiengang Lehramt Mathematik keine Zulassungsbeschränkung wie es sie in den Fächern Biologie und Chemie gibt. Dazu kommt, dass aufgrund der Verpflichtung für ein Hauptfach als Studienfach Mathematik zusätzlich als "zweites Fach" gewählt wird. Dies führt insbesondere noch zu dem Effekt, dass einige Studierende im Lehramt nicht am Propädeutikum teilnehmen, da

sie erst später wissen, dass sie Mathematik studieren werden. Ebenso besitzen, wie beschrieben, die Studierenden in Mathematik, sehr hohe kognitive Grundvoraussetzungen. Diese drei Einflüsse ergeben wesentlich mögliche Erklärungen für die Differenz zwischen den Mathematikstudierenden und den Lehramtsstudierenden. Es zeigt sich hier also in einer noch weiter eingeschränkten Gruppe von Lehramtsstudierenden eine wesentliche Differenz zu den Studierenden des korrespondierenden Studiengangs.
Im Falle der Informatikstudierenden ergibt sich ebenso eine überraschend große Differenz zu den anderen beiden Studiengängen. Diese besitzen in der vorgestellten Studie die höchsten kognitiven Fähigkeiten, auch wenn die dafür kontrollierten mathematischen Fähigkeiten etwas geringer ausfallen als für die anderen Gruppen. Ein hier wesentlicher Unterschied zu der vorgestellten Studie ergibt sich aus dem Profil des Informatikstudiengangs als sehr anwendungsbezogen. Dadurch lässt sich insbesondere davon ausgehen, dass die Studierenden in Informatik verglichen mit anderen Standorten schlechtere Eingangsvoraussetzungen besitzen. Dennoch ist der gefundene Effekt enorm.
Vergleicht man die gezeigten Testergebnisse mit den erwarteten Fähigkeiten im Studium, so ergibt sich für Mathematik und Physik sowie Informatik mit den drei verschiedenen Veranstaltungen des ersten Semesters eine übereinstimmende Reihenfolge. Problematisch hingegen ist der enorme Unterschied der Lehramtsstudierenden gegenüber den Studierenden in Mathematik, mit denen sie gemeinsam die Vorlesungen des ersten Semesters besuchen.
Darüber hinaus ist der Vergleich mit den Schülerinnen und Schüler interessant. So wurde global keine Differenz zu den untersuchten Studienanfängerinnen und -anfängern gefunden, im Vergleich der einzelnen Studiengänge hingegen schon. Dabei ist zu berücksichtigen, dass die Schülerinnen und Schüler eine Positivauswahl darstellen, auch im Bezug auf die Kurse des erhöhten Anforderungsniveaus, da diese freiwillig an der Testung teilgenommen haben. Dass die Studierenden in Mathematik und in Physik bessere Testleistungen erzielen ist dabei dennoch so zu erwarten, da diese beiden Studiengänge gerade die leistungsstärksten Schülerinnen und Schüler in Mathematik erwarten und die höchsten Anforderungen besitzen. Problematischer erscheint es, dass die Studierenden in Informatik und im Lehramt Mathematik schwächere Ergebnisse erzielt haben. In Informatik kann dies durch die oben beschriebene Zusammensetzung und Ausrichtung des Studiengangs erklärt werden. Für die Lehramtsstudierenden ist es hingegen problematisch, wenn diese keine deutlich überdurchschnittlichen Ergebnisse im Mathematiktest erzielen. Dabei ist in diesem Vergleich natürlich die zeitliche Differenz für die Studienanfängerinnen und -anfänger zum Abitur zu bedenken, während die Schülerinnen und Schüler sich gerade in der Vorbereitung auf die Abiturprüfungen befunden haben. Die Bedeutsamkeit dieser zeitlichen Effekte zeigte sich gerade bei der Betrachtung der einzelnen Testgebiete in Abschnitt 7.10. Dennoch zeigen sich die Lehramtsstudierenden in Mathematik in den Ergebnisse nicht

als eine klare Positivauswahl aller Schülerinnen und Schüler. Neben den Problemen, welche dies für das Studium darstellt, ist dies auch problematisch für das spätere Berufsleben, wenn auch fachmathematische Fähigkeiten verlangt werden.

### 9.1.2 Herkunftbundesland - (F3)

Für den Vergleich der vier betrachteten Bundesländer ergaben sich in Abschnitt 7.3 für die Studierenden aus Niedersachsen die geringsten Werte, wobei der Unterschied zu Nordrhein-Westfalen knapp nicht signifikant wurde. Dabei ergaben sich Differenzen von mittlerer Stärke. Die Differenzen zeigten sich dabei nicht zwischen den drei weiteren Bundesländern, sodass hier ein Effekt zwischen den Studierenden aus Niedersachsen und den nicht niedersächsischen Studierenden festzustellen ist.
Wichtig ist zunächst, dass die Ergebnisse keine Evaluationen der jeweiligen Schulsysteme der Bundesländer darstellen. Hier wird stattdessen eine sehr spezifische Gruppe der Studierendenanfängerinnen und -anfänger betrachtet, die starken Auswahleffekten unterliegt. Dabei stellt die Gruppe zunächst einmal nur eine kleine Auswahl der Schulabgängerinnen und -abgänger da, die zusätzliche Entscheidungen bezüglich Studiengang und Studienort getroffen haben.
Daher lassen sich Unterschiede bezüglich der dargestellten Ergebnisse in Abschnitt 3, die sich oft auf die Mittelstufe beziehen erklären, die darüber hinaus auch dort keine großen Effekte zeigten, wenn auch signifikante.
Insgesamt steht das Ergebniss in Einklang mit den beschriebenen Befunden zur räumlichen Nähe zur Universität. Dieser zeigte, dass Studierende, welche für sich selbst eher geringe Erfolge im Studium erwarten, eher in hoher Nähe zum Heimatort studieren. Auch wenn nicht die exakte Entfernung der Heimatorte zu Göttingen gemessen wurde, so stellt der Bundeslandwechsel doch tendenziell eine größere Entfernung dar. Insbesondere zeigt sich die Heimatnähe im Studiengang "Angewandte Informatik", für den knapp 17 % der Studierenden aus Göttingen selbst stammen, in den anderen drei Studiengängen sind es 8 % bis 9 % , diese entspricht dem Durchschnitt der Universität, der 7,7 % beträgt. Für den hohen Anteil der Studierenden mit hoher Heimatnähe ist eher davon auszugehen, dass diese im Schnitt geringe mathematische Fähigkeiten besitzen und die Entscheidung eher auf Grundlage der Ortsentscheidung getroffen haben. Für die Studierenden aus anderen Bundesländern kann davon ausgegangen werden, dass diese sich eher für den Standort entschieden haben.

### 9.1.3 Geschlecht - (F4)

Im Vergleich der beiden Geschlechter zeigt sich in Abschnitt 7.4 ein kleiner Vorsprung in den Testergebnisse für die Studienanfänger gegenüber den Stu-

dienanfängerinnen. Dies entspricht den bisherigen Ergebnissen von kleinen Geschlechtseffekten, der sich auch in der Auswahl auf den Studienanfang übertragen hat.
Gleichzeitig zeigte sich für diese beiden Gruppen eine unterschiedliche Varianz in den Gruppen, wobei die Studienanfänger eine größere Varianz besaßen als die Studienanfängerinnen. Dies führte in der Betrachtung der Spitzengruppe, dem obersten Quartil, zu einer deutlich größeren Differenz zwischen den Geschlechtern, während das untere Quartil für beide Gruppen den gleichen Wert besaß. Dieser Befund entspricht wiederum den vorigen Ergebnissen aus Abschnitt 3.3, dass die Unterschiede zwischen den Geschlechtern sich insbesondere in der stärkeren Spitzengruppe der Jungen finden lässt. Dabei zeigt sich hier der Befund für eine schon vorausgewählte Gruppe des oberen Bereichs.

### 9.1.4 Dauer des Schulbesuchs - (F5)

Abschnitt 7.5 zeigt für die untersuchte Stichprobe einen Effekt mittlerer Stärke der Dauer des Schulbesuchs. Dabei ergeben sich für die Studierenden mit einer kürzeren Schulzeit Vorsprünge von mittlerer Stärke gegenüber den Studierenden mit einer dreizehnjährigen Schulzeit. Dies stellt eine überraschend starke Differenz dar, da in den bisherigen Untersuchungen immer nur kleine Effekte gefunden werden, unabhängig vom genauen Design und auch für alle verschiedenen dargestellten Fächer. Insbesondere gibt es auch in der Richtung des Vorsprungs verschiedene Ergebnisse.
Die Untersuchung, welche die Bundesländer und den Interaktionseffekt beinhaltet, zeigt, dass der Effekt vor allem durch die Studierenden aus Niedersachsen erzeugt wird. Ebenso zeigt sich auch, dass der Effekt schwächer ist als der Effekt des Bundeslandes. Wichtig ist auch hier festzustellen, dass dies keine Evaluation der einzelnen Bundesländer oder der Reform darstellt.
Bei der Betrachtung der Jahre der Umstellung, liegen Niedersachsen und Nordrhein-Westfalen mit dem Jahr 2013 deutlich vor den weiteren Bundesländern. Dies bedeutet insbesondere, dass Studierende aus diesen beiden Bundesländern zum Erhebungszeitpunkt eher einen verzögerten Studienbeginn besitzen, wenn sie ein dreizehnjähriges Abitur absolviert haben. Dies zeigte sich auch in der deskriptiven Stichprobenbeschreibung in Abschnitt 6. Ebenso gab es natürlich auch weitere Entwicklungen des Schulen, welche parallel zu der Umstellung der Schuldauer durchgeführt wurden. Dies könnte zumindest einen Teil des wesentlich größeren Unterschieds in dieser Untersuchung erklären.
Zusammenfassend ergibt sich jedoch in der hier untersuchten Stichprobe ein Vorteil für die Studienanfängerinnen und -anfänger, welche eine kürzere Schulzeit besaßen. Der Vergleich mit den weiteren Einflüssen wird noch einmal in Abschnitt 9.1.7 im Rahmen der gemeinsamen Analyse diskutiert.

### 9.1.5 Verzögerung des Studienbeginns - (F6)

Für die Verzögerung des Studienbeginns zeigt Abschnitt 7.6 einen Effekt von mittlerer Stärke auf das Testergebnis. Dabei ergibt sich zwischen den Studierenden, welche kein Jahr Verzögerung besitzen und den Studierenden, die genau ein Jahr Verzögerung besitzen, eine nur geringe Differenz der durchschnittlichen Testleistung. Dafür zeigen sich für die Studierenden mit mindestens zwei Jahren Verzögerung deutliche Unterschiede verglichen zu den ersten beiden Gruppen. Diese erbrachten auch deutlich geringe Testergebnisse als die Schülerinnen und Schüler.
Diese Ergebnisse entsprechen den Ergebnissen, welche in Abschnitt 3.5 aus den USA vorgestellt wurden. Dabei zeigt hier die Unterscheidung der Verzögerung im Bezug auf ihre Dauer einen wesentlichen Effekt. Die Studienanfängerinnen und -anfänger mit genau einem Jahr Verzögerung besitzen zwar schwächere Testergebnisse als die direkt anfangenden Studierenden, diese Differenzen sind aber gering. Dass sie danach wesentlich größer werden, lkönnte bedeuten, dass ein Teil des Effekts auch darauf zurückzuführen ist, dass die schon vorher schwächeren Schülerinnen und Schüler ihre Studienaufnahme verzögern. Denn es stellt sich die Frage, ob schwäche Schülerinnen und Schüler eher verzögert anfangen zu studieren oder sie im Laufe der Verzögerung mathematische Inhalte vergessen. Diese Frage lässt sich mit den hier erhobenen Daten global nicht beantworten, da dafür Daten während der Wartezeit erhoben werden müssten. Hinweise darauf werden in Abschnitt 9.3 über die einzelnen Bereich beschrieben.
Insgesamt lässt sich dennoch festhalten, dass der verzögerte Studienbeginn mit schwächeren Testleistungen einhergeht, insbesondere ab dem zweiten Jahr.

### 9.1.6 Kursform - (F7)

Abschnitt 7.7 zeigt im Vergleich der Kursformen "erhöhtes Anforderungsniveau" und "grundlegendes Anforderungsniveau" einen großen Effekt in den Testleistungen. Dies entspricht den bisherigen Ergebnissen wie sie in Abschnitt 3.7 beschrieben wurden. Dies stellt eine enorme Differenz für die Studienanfängerinnen und -anfänger in ihren Eingangsvoraussetzungen dar, und bedeutet den größten gefundenen Einzeleffekt auf einem Niveau mit dem gewählten Studiengang. Daher bedeutet die Kurswahl aber auch ein großes Risiko für Studierende insbesondere in Mathematik und Physik, welche in der Schule einen Kurs auf grundlegendem Anforderungsniveau besucht haben, da zu erwarten ist, dass sie geringe Kenntnisse als Eingangsvoraussetzung besitzen. Die Kurswahlentscheidung in der Schule zeigt hier also bereits einen wesentlichen Einfluss auf spätere Studienentscheidungen. Dabei ist hier nicht zu klären, ob die Kurswahl eher eine Folge schwächere mathematischer Fähigkeiten darstellt oder sie schwäche Fähigkeiten besitzen, weil sie eine anderen Kurs gewählt haben. Auch hier gibt die spätere Betrachtung der einzelnen Gebiete einige Hinweise. Es bedeutet die Wahl eines Kurses auf grundlegendem Anforderungsniveau also ein erhöhtes Risiko für die spätere Wahl eines Studiums in Mathematik oder Physik.

### 9.1.7 Übersicht aller Prädiktoren - (F8)

In der Analyse aller nichtmetrischen Prädiktoren zeigt Abschnitt 7.8, dass alle oben beschriebenen Variablen weiterhin einen signifikanten Einfluss auf das Testergebnis besitzen, der Erhebungszeitraum jedoch wiederum nicht. Dabei erklärt das Modell insgesamt 40,8 % der gesamten Varianz in der Stichprobe. Den größten Einfluss besitzt dabei die Variable Studiengang, die knapp 15 % Varianzaufklärung besitzt, wenn die anderen Einflüsse nicht betrachtet werden. Als weitere wesentliche Variablen werden die Kursform, das Bundesland sowie die Schuldauer aufgenommen. Die Verzögerung des Studienbeginns sowie das Geschlecht liefern weitere geringere Beiträge.
Dass der Studiengang den höchsten Beitrag liefert entspricht dabei in Teilen den Erwartungen. Da die Studiengänge fachspezifisch gewählt werden und verschiedene Anteile an Mathematik beinhalten, kann erwartet werden, dass sich in dieser Fähigkeit klare Unterschiede zeigen. Dabei ist auf der anderen Seite zu bedenken, dass die hier betrachteten Studiengänge bereits eine Einschränkung auf eine Gruppe ähnlicher Studiengänge darstellen. Dies weist darauf hin, dass die Entscheidung für einen Studiengang eng mit den Fähigkeiten verknüpft sein könnte und wesentlicher als die Variablen zum Schulhintergrund, die erhoben wurden.
Auch der noch hohe Beitrag der Kursform ist nicht überraschend, da die Kurswahl ja auch eine Folge von erbrachten Leistungen darstellt sowie weiter die Unterschiede in der Sekundastufe II vergrößern sollte. Dennoch geht dieser Beitrag im Vergleich mit den anderen Prädiktoren auf hohem Niveau etwas zurück. Hier zeigt sich dann wiederum, dass die Schuldauer sowie das Bundesland einen nicht geringen Einfluss besitzen. Insbesondere für die Schuldauer bekräftigt dies die überraschend hohe Differenz zwischen den beiden Schuldauern. Hier war die vorige Interpretation aus Abschnitt 9.1.4 wegen der Konfundierung mit der Verzögerung des Studienanfangs nicht eindeutig. Im gemeinsamen Vergleich hier zeigt sich aber, dass der Effekt auf die Schuldauer zurückführbar ist. Damit ergibt sich also insgesamt in dieser Stichprobe der Vorteil Studierender mit geringerer Schulzeit. Der Effekt der zeitlichen Verzögerung des Studienbeginns verschwindet in der gemeinsamen Analyse.
Das Geschlecht leistet weiterhin nur einen geringen Beitrag für das Testergebnis. Allerdings zeigt sich dieser als unabhängig von den anderen untersuchten Variablen. Insbesondere zeigen sich die Unterschiede auch bei Kontrolle der Interessenswahlen von Studiengang sowie der Kursform.
Insgesamt lassen sich gut 40 % der Varianz in der Testleistung durch die be-

trachteten Variablen erklären. Dabei werden hier die beiden Schulzensuren explizit nicht betrachtet. Damit ist diese Varianzaufklärung überraschend hoch. Der große Anteil des Studiengangs stellt dabei noch einmal deutlich heraus wie deutlich sich die Eingangsvoraussetzungen zwischen den Studiengängen unterscheiden. Der Beitrag der Schulzeit ist hier überraschend deutlich, ebenso der Beitrag des Bundeslands, welcher sich auf die Spezifität des Einzugsgebiets der Universität Göttingen zurückführen lässt.

### 9.1.8 Schulleistungen - (F9/10)

Für die Korrelationen zwischen Testergebnissen und Schulleistungen zeigt sich in Abschnitt 7.7 mit Korrelationen knapp unter 0,6 auch ein Ergebnis, welches den beschriebenen früheren Studien aus Abschnitt 3.7 entspricht. Die Korrelation liegt allerdings am oberen Rand der vorigen Ergebnisse. Insbesondere auch der höhere Zusammenhang für Schülerinnen und Schüler aus dem Kurs mit erhöhtem Anforderungsniveau bestätigt sich hier. Dabei bleiben aber auch in Kursen mit grundlegendem Anforderungsniveau relativ hohe Zusammenhänge bestehen. Dies ist zum Teil auf die Beschränkung auf Studierende zurückzuführen, da hier viele unsystematische Effekte am unteren Leistungsspektrum herausfallen sollten. Dies gilt ebenso für die Höhe in Bezug auf den Kurs mit erhöhtem Anforderungsniveau, da sich auch hier nur Teilnehmende mit einer weiteren Entscheidung für einen mathematischen Studiengang in der Stichprobe befinden. Eine Differenz zwischen Abiturnote und der letzten Mathenote im Bezug zum Testergebnis lässt sich hier nicht feststellen. Dies zeigt sich insbesondere auch am hohen Zusammenhang zwischen letzter Mathematiknote und Abiturnote. Die fachspezifische Schulnote zeigt hier also einen höheren Zusammenhang als in den beschriebenen früheren Studien.
Die Betrachtung der Regressionsanalyse in Tabelle 16 zeigt, dass mit einem Bestimmtheitsmaß von $R^2 = 0,48$ knapp die Hälfte der Varianz im Testergebnis durch die drei leistungsbezogenen Prädiktoren erklärt werden kann. Dabei tragen alle drei Prädiktoren einen signifikanten Beitrag zur Aufklärung bei. Der Vergleich der standardisierten Koeffizienten zeigt dabei, dass diese die gleiche Größe besitzen, die Effekte der drei Prädiktoren also die gleiche Größe besitzen.
Der Vergleich der unstandardisierten Koeffizienten zeigt hier einen Unterschied zwischen den beiden Kursformen von 0,6 unter Kontrolle der beiden anderen Variablen; der Effekt bei alleiniger Betrachtung der Kursform wird also wesentlich vermindert. Die Betrachtung der unstandardisierten Koeffizienten zeigt ebenso, dass eine Verschlechterung in der Abiturnote um eine Note mit einer Verschlechterung im Testergebnis von 0,438 einhergeht. Der Einfluss der Abiturnote ist also nicht nur eng, sondern auch bedeutsam. Ebenso ergibt sich für die letzte Mathenote eine Verbesserung um 0,303 in der Testleistung bei einer Verbesserung um drei Notenpunkte, also eine No-

te. Der Einfluss liegt also etwas geringer, aber bei einer Standardabweichung von knapp 0,9 in den Testergebnisse immer noch wesentlich. Insgesamt bestätigt sich also auch hier der hohe Zusammenhang zwischen Leistungen in der Schule und den mathematischen Fähigkeiten. Insbesondere zeigen sich noch deutliche fachspezifische Einflüsse über die allgemeine Abiturnote hinaus.

### 9.1.9 Lineare Regression mit allen Prädiktoren - (F10)

In der Regressionsanalyse in Abschnitt 7.9 mit allen Prädiktoren werden neben den Faktoren aus der vorigen Analyse auch die Beiträge zum Geschlecht, der Schuldauer, dem Herkunftsbundesland und dem Studiengang signifikant, jeweils für die Gruppen, die auch zuvor signifikante Beiträge lieferten. Dabei zeigt sie Regression aus dem vorigen Abschnitt auch die ersten drei größten Beiträge, welche weiter ergänzt werden. Dabei liefern die zusätzlichen Beiträge weitere 10 % Varianzaufklärung. Damit bedeuten diese zusätzlichen Beiträge einen bedeutsamen Unterschied für die Testleistung, welcher über die Leistungen aus der Schule hinausgeht.

Der Beitrag der letzten Mathematiknote geht dabei gegenüber der Abiturnote zurück. Dies entspricht den bisherigen Ergebnissen. Es ergibt sich also bei der gemeinsamen Betrachtung der Variablen der höhere Zusammenhang zwischen Abiturnote und mathematischer Fähigkeiten als für die Einzelnote.

Dabei zeigt sich für den ersten Beitrag, das Geschlecht, eine Vergrößerung des Beitrags unter Kontrolle der weiteren Prädiktoren. Die Dauer des Schulbesuchs hingegen verliert etwas an Unterschied. Ebenso leichte Verschiebungen gibt es für die verschiedenen Herkunftsbundesländer. Diese Beiträge zeigen sich also wesentlich unverändert von der Kontrolle der Schulleistungen.

Der Unterschied zwischen den Geschlechtern bleibt hier unter Kontrolle der weiteren Variablen ziemlich konstant, ist also unabhängig von den anderen Variablen. Dabei vergrößert sich sogar der Abstand der Gruppen im Vergleich zum einfaktoriellen Vergleich.

Die Differenz zwischen den verschiedenen Studiengängen hingegen geht deutlich zurück und liefert bei Betrachtung der standardisierten Koeffizienten die geringsten Beiträge. Die Differenz zwischen den Studiengängen lässt sich also zu einem großen Anteil auf die Schulleistungen zurückführen. Darüber hinaus zeigen sich dennoch unter deren Kontrolle Nachteile in den Testleistungen für die Studierenden im Lehramt sowie in Informatik. Dabei sind diese auch in ihrer Höhe bedeutsam.

Insgesamt lässt sich die Varianzaufklärung von knapp 60 % als hoch ansehen, da an dieser Stelle keine nicht-kognitiven Persönlichkeitsmerkmale aufgenommen wurden. Dabei bestätigte sich insgesamt der hohe Zusammenhang zwischen Abiturnote und mathematischen Fähigkeiten. Dennoch wurden noch 10 % der Varianz darüber hinaus durch die weiteren Variablen aufgeklärt, die jeweils weitere Beiträge lieferten. Nur die verzögerte Studienaufnahme verschwand unter Kontrolle der weiteren Variablen.

## 9.2 Klausurergebnisse

Die Untersuchung auf die Klausurergebnisse am Ende des ersten Semesters stellt sich dreischrittig dar. Die erste Frage ist, wer überhaupt an einer Klausur teilnimmt, danach folgt die Frage, ob die Klausur bestanden wird, und am Ende die Zensur, mit der bestanden wird. Entsprechend diesem Vorgehen sollen nun die Ergebnisse diskutiert werden.

### 9.2.1 Klausurteilnehmende - (K1/K2)

Abschnitt 8.2 zeigt, dass ungefähr 70 % der Teilnehmenden des Propädeutikums an einer Klausur am Ende des ersten Semesters teilnehmen. Dies bedeutet aber auch, dass von den Anfängerinnen und Anfängern bereits ein Drittel an keiner Klausur teilnimmt. Dabei ist nicht auszuschließen, dass zu einem späteren Zeitpunkt noch eine dieser Klausuren angetreten wird.
Dennoch bedeutet dies für eine ziemlich große Zahl der Studierende, dass sie im ersten Semester keine zählbare Leistung erbringen. Hierbei zeigt sich aufgrund der Differenz im Testergebnis der Teilnehmenden und den Nicht-Teilnehmenden an einer Klausur die mathematischen Kenntnisse einen wesentlichen Einfluss haben. Dies wiederum bestätigt den wesentlichen Einfluss der Vorkenntnisse zu Studienbeginn auf das erste Semester. Dabei wird hier sogar nur das Teilnehmen ohne Ergebnisbetrachtung untersucht.

### 9.2.2 Klausurbestehen - (K3)

Im Vergleich der beiden untersuchten Jahrgänge verändert sich die Zusammensetzung der Studierenden, welche an der Klausur teilgenommen haben, da sie für die Physik-Studierenden keine Pflichtveranstaltung mehr darstellte. Die Untersuchungen in Abschnitt 8.3 bezüglich den Veränderungen zwischen den beiden Jahren ergeben dabei deutlich, dass die Unterschiede zwischen den beiden Jahren nicht auf diese Differenz zurückzuführen sind. Das Ausschließen der Physikstudierenden sowie einzelne Berechnungen der Modelle ergeben keine Verbesserungen der Prädiktion. Ebenso wird von vornherein der Studiengang nicht in die Modelle aufgenommen. Dies weist also eher darauf hin, dass der Effekt des Jahrgang, welcher sich für "Differential- und Integralrechnung" zeigt eher auf spezifische Eigenschaften der Veranstaltungen zurückführbar ist als auf die veränderte Zusammensetzung der Studierenden. Insbesondere ergibt sich, dass eine gemeinsame Modellierung der beiden Jahrgängen möglich ist, auch wenn es im einzelnen Unterschiede zwischen den Veranstaltungen gibt.
Insgesamt zeigt sich, dass die Aufklärung für "Differential- und Integralrechnung" etwas besser gelingt als für "Analytische Geometrie und lineare Algebra" , wobei ungefähr 80 % der Fälle mit den Modellen korrekt zugeordnet werden. Dies stellt eine wesentliche Verbesserung gegenüber dem Modell dar, welches keine Variablen beinhaltet und jeweils ungefähr 60 % der Fälle

(alle fallen durch) zuordnet. Dabei wurden in keinem Fall nichtleistungsbezogene Variablen in die Modelle aufgenommen. Damit werden insbesondere kein Geschlecht und kein Studiengang benachteiligt.
In das jeweils beste Modell wird das Testergebnis aufgenommen. Dabei ergibt sich hieraus insbesondere für die Veranstaltung "Differential- und Integralrechnung" eine Verbesserung gegenüber der Abiturnote. Insgesamt bestätigt sich damit, dass die mathematischen Fähigkeiten, wenig überraschend, die zentrale Rolle für den Erfolg im mathematischen Studium darstellen. Dabei ist eine korrekte Zuordnung von 80 % der Studierenden zwar hoch, auf der anderen Seite zeigt sich auch, dass innerhalb des ersten Semesters noch weitere Merkmale einen wesentlichen Beitrag leisten.

### 9.2.3 Klausurzensuren

Für die Betrachtung der Zensuren im ersten Semester zeigt sich in Abschnitt 8.4 das Testergebnis als bester Prädiktor im Vergleich der drei Leistungsmaße. Dies bestätigt sich für beide Lehrveranstaltungen. Mit Korrelationen von ungefähr r=0,6 sind dabei die Zusammenhänge höher als die Abiturnote oder die letzte Mathematiknote. Die beiden Leistungsdaten aus der Schule liegen dabei in den auch vorher für Universitätsleistungen bekannten Größenordnung. Der Mathematiktest zeigt hingegen eine wesentlich höhere Korrelation als für die vorher beschrieben Studien. Dabei ist zu bedenken, dass dieser Test spezifisch für das erste Semester und einen Standort entwickelt wurde und dementsprechend einen hohen Zusammenhang besitzen sollte. Dieses Ergebnis bedeutet dabei allerdings, dass dies gelungen ist.
Die linearen Regressionen der Klausurergebnisse bestätigen hier das Ergebnis aus dem vorigen Abschnitt, dass über die leistungsbezogenen Variablen hinaus keine der erhobenen Variablen einen signifikanten Beitrag leistet. Insbesondere besitzen damit das Geschlecht und das Herkunftsbundesland keinen Einfluss. Interessant ist auch der Befund, dass es keine Unterschiede zwischen den Studiengängen gibt, wenn die Vorleistungen kontrolliert werden. Hier finden also auch keine Vergrößerungen von Differenzen statt.
Es bestätigt sich hier nicht der vorige Befund mit Daten aus dem Jahr 2012 (Halverscheid & Pustelnik, 2013) des Einflusses von Schuldauer und Verzögerung auf Klausurergebnisse. Im Vergleich der beiden Vorlesungen zeigt sich hier jedoch wiederum der größere Zusammenhang bzw. der größere Effekt des Vorwissenstests auf die Zensur in "Differential- und Integralrechnung" als auf "Analytische Geometrie und lineare Algebra". Insgesamt ergibt sich also über drei Jahre ein ähnliches Bild, sowohl für den Vergleich der beiden Veranstaltungen als auch für die einzelnen Prädiktoren.
Werden die beiden verschiedenen Vorlesungen verglichen, zeigen sich grundlegend, insbesondere mit dem Testergebnis, ähnliche Regressionsgleichungen. Im Speziellen gibt es hingegen schon Unterschiede. Zunächst wird für die erste Klausur in "Analytische Geometrie und lineare Algebra" die letz-

te Mathematiknote aufgenommen, auch vor der Abiturnote. Insbesondere ist auch der Zusammenhang insgesamt schwächer als für die "Differential- und Integralrechnung" zu den Testergebnisse. So liegt der unstandardisierte Koeffizient um ungefähr 0,25 niedriger. Das bedeutet eine Verschlechterung im Testergebnis geht mit einer geringen Verschlechterung in den Klausurergebnisse einher. Dies könnte damit zusammenhängen, dass die Inhalte der "Differential- und Integralrechnung" näher an den Schulinhalten liegen und diese vertiefen als für die "Analytische Geometrie und lineare Algebra", in der neue Strukturen eingeführt werden.
Im Vergleich der ersten Klausurnote und der endgültigen Zensur, zeigt sich in "Analytische Geometrie und lineare Algebra" keine wesentliche Veränderung, während für "Differential- und Integralrechnung" der Zusammenhang wesentlich zurückgeht. Dies stellt einen Hinweis darauf dar, dass zwischen den Klausurterminen in den Semesterferien nach den Vorlesungen sich noch einiges verändern lässt. Dabei ist der Unterschied zwischen den beiden Veranstaltungen nicht aufklärbar.
Die Unterschiede in den Anfangsjahren lassen sich wohl auch hier wesentlich auf die Dozenten der Veranstaltungen zurückführen. Der Vergleich wurde im vorigen Abschnitt, insbesondere bezogen auf die veränderte Studierendenzusammensetzung, beschrieben.
Insgesamt zeigt sich also, dass der Test eine hohe Prädiktionskraft für das Klausurergebnis im ersten Semester erzielt und dahingehend auch über der Abiturnote sowie der letzten Mathematiknote liegt. Dies entspricht den bisherigen Ergebnissen. Dabei zeigen für diese Auswertungen die Betrachtungen einzelner Testgebiete keine Verbesserung in den Prädiktionen für die Klausurergebnisse.

## 9.3 Teilgebiete - (F11/K5)

Für die verschiedenen Teilgebiete ergeben sich in Abschnitt 7.10 zunächst in den Varianzaufklärungen deutliche Unterschiede. Es zeigt sich für alle Gebiete, dass diese geringer ausfällt als für den Gesamttest. Dabei gibt es eine Tendenz zu höherer Aufklärung für die in der Schule späteren Inhalte. Dies bedeutet, dass für die in der Schule früher behandelten Gebiete andere Faktoren einen größeren Einfluss besitzen, welche hier nicht betrachtet wurden. Die weiteren Variablen können dabei einen systematischen Einfluss besitzen oder einfach aufgrund der zurückliegenden Zeit verschiedene Effekte. Der Zusammenhang könnte auch plausibel erklärt werden mit Effekten des Vergessen von zurückliegenden Inhalten. Hierfür würden sich im Laufe der Zeit unsystematische Unterschiede ergeben. Für die späteren Gebiete zeigen sich diese dann weniger.
Im verschiedenen Einfluss der untersuchten Variablen auf verschiedene Ergebnisse zeigten sich interessante Ergebnisse für die Kursform sowie die zeitliche Verzögerung. Dabei ergaben sich für die Kursform klare Differenzen

zwischen den Gebieten vor und nach dem Beginn der Oberstufe. Dies zeigt, dass die Unterschiede nach der Kurswahl noch einmal zunehmen und sich nicht nur aus den Unterschieden der vorherigen Schulzeit ergeben.
Um zeitliche Effekte genauer zu untersuchen, eignet sich die Betrachtung der zeitlichen Verzögerung des Studienbeginns, welche sich neben einer systematischen Auswahl im Bezug auf den Studienbeginn auch zeitlich auswirkt. Dabei wurden, wie in Abschnitt 3.5 dargestellt, explizit nichtleistungsbezogene Gründe als Hauptursachen der Verzögerung angegeben. Die Betrachtung der zeitlich verzögerten Studienaufnahme zeigt insgesamt nur einen kleinen Effekt, aber auch hier lässt sich die Tendenz der erhöhten Varianzaufklärung für die späteren Gebiete finden, wie sonst nur bei der Kursform.
Die Betrachtung der Mittelwertunterschiede spiegelt diese Tendenz auch wider. Insbesondere passen hier auch die Ergebnisse der untersuchten Schülerinnen und Schüler sehr gut in das Bild. So nehmen insgesamt alle Differenzen der Mittelwerte mit dem zeitlichen Verlauf des Gebiets in der Schule zu. Insbesondere gibt es in den frühen Gebieten keine Unterschiede zwischen den Schülerinnen und Schülern zu den am längsten verzögerten Studierenden, während sich für die späteren Gebiete sogar starke Effekte zeigen.
Dies deutet auf einen Effekt des Vergessens bzw. Konsolidieren des Wissens hin: Für die in der Schule früh behandelten Gebiete ändert sich mit dem zeitlichen Verlauf später eher wenig, da die Entwicklung des Vergessens oder Konsolidierens bereits abgeschlossen ist, in den später behandelten Gebieten hingegen spielt die Verzögerung eine größere Rolle. Der genaue Mechanismus als Ursache der Ergebnisse ist natürlich mit dem hier durchgeführten Design nicht zu entscheiden, dennoch ist die Konsistenz der Ergebnisse, auch über die Gruppen hinweg, hoch.
In Bezug auf die Prädiktion von Klausurerfolg zeigte sich die Aufteilung der verschiedenen Gebiete als weniger hilfreich. Hier stellte das Gesamttestergebnis die wesentlich bessere Prädiktion dar. Insbesondere zeigten sich auch keine zeitlichen Effekte oder Effekte der Passung zu den Vorlesungen. Damit ergeben sich insgesamt also spezifische Einflüsse auf die Leistungen in den verschiedenen Testgebieten zu Beginn des Studiums. Auf die Ergebnisse am Ende des ersten Semesters ist die Wirkung allerdings eher gering. Es ist dabei jedoch zu bemerken, dass als Teil des Propädeutikums Studierenden mit spezifischen Schwächen in einzelnen Teilgebieten gezielt in diesen Gebieten gefördert wurden. Es zeigten sich dabei zum Teil große Aufholeffekte (Halverscheid, Pustelnik, Schneider & Taake, 2014). Mögliche Effekte auf das Studium wurden also systematisch verkleinert.

# 10 Kritik und Ausblick

Die Arbeit gibt für einen Universitätsstandort und vier ausgewählte Studiengänge eine genaue Darstellung der fachmathematischen Eingangsvoraussetzungen der Studierenden sowie längsschnittliche Daten für den Klausurerfolg am Ende des ersten Semesters an. Dabei wurde ein Testinstrument verwendet, welches am oberen Rand der Erwartungen der Bildungsstandards bzw. des niedersächsischen Kerncurriculums liegt. Dementsprechend ist eine Verwendung für weitere Studiengänge nicht gewährleistet. Denn die Auswahl der vier Studiengänge stellt eine Auswahl der Studiengänge dar, für welche ein sehr hoher Mathematikanteil mit entsprechenden Voraussetzungen zu erwarten ist. Dennoch ermöglicht gerade diese Beschränkung auch spezifische Vergleiche. Dabei ergaben sich, insbesondere zwischen den Studiengängen, doch enorme Unterschiede. Es zeigt sich also, dass auch eine fokussiertere Betrachtung auf eine spezifische Studierendengruppe lohnend ist.

Neben der Auswahl spezifischer Studiengänge wurde auch nur ein Universitätsstandort betrachtet. Dabei ist davon auszugehen, dass sich die Ergebnisse für andere Universitäten unterscheiden können. Dies gilt insbesondere für Effekte der Bundeslandauswahl. Die Fokussierung ermöglicht auf der anderen Seite eine klarere Beschreibung der Situation an dem einen Standort. Eine Verbreiterung der Erhebung an weiteren Standorten ergäbe dann die Möglichkeit, die Voraussetzungen von Studienanfängerinnen und -anfängern zu vergleichen. Insbesondere kann davon ausgegangen werden, dass sich die beschriebenen Einflussfaktoren, abgesehen von der dominanten Rolle der Abiturnote, in ihrer Gewichtung unterscheiden, was auf Profile und Einzugsgebiete der Universitäten zurückzuführen ist. Entsprechend bietet ein Vergleich dann die Chance, Spezifika der einzelnen Standort herauszuarbeiten und davon abhängig verschiedene Folgen abzuleiten.

Die domänenspezifische Betrachtung der Testergebnisse für die Eingangsvoraussetzungen zeigt, dass es hier Unterschiede zwischen mathematischen Bereichen gibt. Auch wenn insgesamt der verwendete Test als eindimensional skaliert werden konnte, zeigen sich Unterschiede zwischen den vorher inhaltlich definierten Bereichen. Dabei ergeben sich auch kleine Effekte, welche in einen zeitlichen Verlauf eingeordnet werden konnten. Insbesondere die Einflussfaktoren sowie die Varianzaufklärung unterschieden sich wesentlich für die verschiedenen Gebiete. Trotz der Eindimensionalität lohnt sich also auch hier eine Differenzierung in Bezug auf die verschiedenen mathematischen Gebiete.

In Bezug auf das Vorwissen ergibt sich also insgesamt eine Einschränkung auf die untersuchten Studierenden, auf eine Universität und wenige Studiengänge. Dennoch ergeben sich deutliche Unterschiede auch innerhalb dieses fokussierten Blicks. Auf der anderen Seite ergibt die Ausdifferenzierung der mathematischen Gebiete dennoch zusätzlichen Erkenntnisgewinn.

Für die Betrachtung der Klausurergebnisse zeigt sich, dass neben den Lei-

stungsvariablen, die weiteren erhobenen Variablen keine signifikanten Beiträge leisten. Dies entspricht den vorgestellten früheren Ergebnissen. Dabei ist zu beachten, dass weitere nichtkognitive Persönlichkeitseigenschaften, wie zum Beispiel Interesse und Motivation, nicht betrachtet werden. Auch ohne diese Variablen wurden dabei korrekte Zuordnungen von bis zu 85 % erreicht, sodass wesentliche Verbesserungen nur über diesen Wert hinaus zu erzielen sind. Zusätzlich zu solchen Erklärungsvariablen lassen sich auch weitere Zielvariablen untersuchen, wie Zufriedenheit, welche auf den Abbruch auch wesentlichen Einfluss besitzen könnten, wobei auch hier Leistungsschwierigkeiten als zentrales Kriteriums des Abbruchs aus früheren Studien bereits bekannt sind. Dennoch ist davon auszugehen, dass gerade die nichtkognitiven Variablen in der konkreten Abbruchsentscheidung bzw. Bleibeentscheidung eine wesentliche Rolle spielen werden.
Auch für die Klausurergebnisse ist eine Ausweitung auf weitere Universitäten interessant. Da hier der Test standortspezifisch, durch die Vorlesungsskripte, entwickelt wurde, müssen Ergebnisse nicht übertragbar sein. Die Validität ist aufgrund der Konstruktion aus dem niedersächsischen Kerncurriculum zwar zumindest für Niedersachsen im Bezug auf das Schulwissen standortunabhängig, für den Hochschulbezug hingegen nicht. Es könnten außerdem für die Regressionsmodelle, wie auch bei den Eingangsvoraussetzungen, deutliche standortabhängige Effekte erwartet werden. Für die Universität Göttingen zeigt sich jedoch, auch wenn es Unterschiede in den Bestehensquoten und auch Zensuren gibt, dass eine gemeinsame Modellierung über die Jahrgänge möglich ist. Dabei zeigt sich insbesondere das Verhältnis der beiden Veranstaltungen über die Jahre hinweg als konstant.
Eine einzelne Betrachtung der beiden Veranstaltungen des ersten Semesters ergibt einen klaren Unterschied im Zusammenhang zu den Eingangsvoraussetzungen und den Klausurergebnissen. Diese lagen für die Veranstaltung "Differential- und Integralrechnung" höher als für die Veranstaltung "Analytische Geometrie und lineare Algebra". Dies könnte auf den höheren Anteil von Analysis in der Schule zurückgeführt werden. Die Aufteilung der einzelnen Gebiete ergibt für die Klausurergebnisse, anders als für die Eingangsvoraussetzungen, jedoch keinen zusätzlichen Informationsgewinn. Für die Berachtung des Studienerfolgs im Sinne von Klausurbestehen ergibt sich also eine Abhängigkeit von der gewählten Veranstaltung als Zielvariablen. Es zeigt sich insbesondere, dass aufgrund des Vorwissens zu Beginn des Studiums, welche ausschließlich auf Schulwissen bezogen war, gute Vorhersagen für das Ende des ersten Semesters möglich sind, trotz der beschriebenen Brüche zwischen Schulmathematik und Universitätsmathematik.
Neben der Beschreibung der Studierenden des ersten Semesters wurde auch ein Vergleich mit den Schülerinnen und Schülern kurz vor ihrem Abitur durchgeführt. Dabei ist zu beachten, dass diese Vergleichsgruppe einem klaren Auswahleffekt unterliegt, da die Schülerstichprobe aus Freiwilligen besteht. Dabei ist eher von einer Positivauswahl auszugehen. Die Einordnung

der Schülerinnen und Schüler ist dabei vor allem im Bezug auf die Studienfachwahl und zugehörigen Leistungen interessant. Hier schneiden die Lehramtsstudierenden schwächer ab als die untersuchten Schülerinnen und Schüler. Diese Einordnung unterstreicht dabei die Notwendigkeit auch hier die verschiedenen Studiengänge auszudifferenzieren. Zur Einordnung der gefundenen Ergebnisse sei dennoch noch einmal auf den Auswahleffekt hinzuweisen, sodass ein umfassenderer Vergleich auf Seiten der Schülerinnen und Schüler wünschenswert ist, um entsprechende Einordnungen vorzunehmen.
Eine Einschränkung der Studienergebnisse stellt die Ziehung der Stichprobe dar. Es wurden ungefähr 30 % der Studienanfängerinnen und -anfänger untersucht. Da diese am ersten Tag des Propädeutikums erhoben wurde, können hierbei systematische Verzerrungen stattfinden. Dafür wurden im Propädeutikum alle Teilnehmenden befragt. Dabei ist nicht ganz klar, in welche Richtung sich diese Verzerrungen auswirken. Auf der einen Seite könnten gerade die Studierenden zum Propädeutikum kommen, welche sich vor dem Studium besonders informiert haben und entsprechend eine Positivauswahl darstellen. Auf der anderen Seite könnten gerade die Studierenden, welche besonders gut sind, sich gegen die Teilnahme entscheiden, in dem Glauben dieses nicht nötig zu haben, sodass eher die Schwächeren das Propädeutikum besuchen. Eine besondere Gruppe stellen dabei die Lehramtsstudierende dar. Diese können zum Teil zu Beginn des Propädeutikums noch nicht wissen, dass sie Mathematik studieren. Dies bedeutet tendenziell auch eher schwäche Fähigkeiten in Mathematik. Insgesamt lässt sich also nicht genau klären, ob und ich welche Richtung die mathematischen Eingangsfähigkeiten der Stichprobe sich von allen anderen Studienanfängerinnen und -anfängern unterscheiden. Insbesondere lässt dies keine Aussagen über die Folgen der Teilnahme oder Nicht-Teilnahme am Propädeutikum zu, auch wenn es in Bestehensquoten und Durchschnittsnoten keine Unterschiede gibt.
Nicht Teil der Untersuchung stellen der genaue Zeitpunkt des Studienabbruchs und der zugehörige Prozess dar. Da nur 70 % der Studienanfängerinnen und -anfänger überhaupt am Ende des ersten Semesters an einer Klausur teilnehmen, stellt sich die Frage, was mit den übrigen 30 % der Studienanfängerinnen geschieht. Hierbei ist zunächst unklar, ob diese 30 % niemals eine Klausur schreiben werden oder nur am Ende des ersten Semesters keine Klausur schreiben. Zusätzlich ist zunächst deskriptiv von Interesse, wann genau im ersten Semester der Abbruch stattfindet, dies wäre mit Hilfe von Übungszettelbearbeitungen möglich. Hierbei sind dann auch die weiteren, hier nicht betrachteten Variablen wie Zufriedenheit und Interesse, zu untersuchen. Darüber hinaus könnte sich hier konkreter der Einfluss domänenspezifischer Stärken und Schwächen von Studierenden zeigen, bzw. zeigen, wo das erhobene Vorwissen genau zum Einsatz kommt.
Neben der Zeit bis zu der ersten Klausur ist auch die Zeit danach weiter interessant. Wie die Unterschiede zwischen erster und letzter Klausur gezeigt haben, ergeben sich hier noch wesentliche Änderungen. Daher ist die Reakti-

on auf einen Nicht-Bestehen untersuchenswert. Hier kommen auch die nicht vernachlässigbaren Abbruchquoten nach dem ersten Semester bzw. Jahr ins Spiel.

# 11 Literatur


Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotugu Akaike* (S. 199–213). Springer.

Andrietti, V. (2016). The causal effects of an intensified curriculum on cognitive skills: Evidence from a natural experiment. *Economic Series 16-06*.

Andrietti, V. & Su, X. (2018). The impact of schooling intensity on student learning: Evidence from a quasi-experiment. *Education Finance and Policy*, 1–39.

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2015). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung.* Springer-Verlag.

Baumert, J., Bos, W. & Lehmann, R. (2000). *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (Bd. 2). Springer-Verlag.

Bausch, I., Biehler, R., Bruder, R., Fischer, P. R., Hochmuth, R., Koepf, W., … Wassong, T. (2014). *Mathematische Vor-und Brückenkurse Konzepte, Probleme und Perspektiven.* Springer.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289–300.

Berger, M. (2004). The functional use of a mathematical sign. *Educational Studies in Mathematics*, *55* (1-3), 81–102.

Bloch, I. & Ghedamsi, I. (2004). The teaching of calculus at the transition between upper secondary school and the university: Factors of rupture. In V. Durand-Guerrier, S. Soury-Lavergne & F. Arzarello (Hrsg.), *Proceedings of the sixth congress of the european society for research in mathematics education* (Bd. 12). Citeseer.

Bozick, R. & DeLuca, S. (2005). Better late than never? Delayed enrollment in the high school to college transition. *Social Forces*, *84* (1), 531–554.

Büttner, B. & Thomsen, S. L. (2015). Are we spending too many years in school? Causal evidence of the impact of shortening secondary school duration. *German Economic Review*, *16* (1), 65–86.

Clark, M. & Lovric, M. (2008). Suggestion for a theoretical model for secondary-tertiary transition in mathematics. *Mathematics Education Research Journal*, *20* (2), 25–37.

Clark, M. & Lovric, M. (2009). Understanding secondary–tertiary transition in mathematics. *International Journal of Mathematical Education in Science and Technology*, *40* (6), 755–776.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and psychological measurement*, *33* (1), 107–112.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112* (1), 155.

Derr, K. & Hübl, R. (2014). Mathematik Studienvorbereitung Online für technische Studiengänge. *Qualifizierung von Studierenden im Student-Life-Cycle*.

Derr, K., Hübl, R. & Ahmed, Z. (2013). Online-Eingangstests und Lernmaterialien zur Studienvorbereitung Mathematik in den Ingenieurwissenschaften. *Mathematik im Übergang Schule/Hochschule und im ersten Studienjahr*, 35.

Dieter, M. (2012). *Studienabbruch und Studienfachwechsel in der Mathematik: Quantitative Bezifferung und empirische Untersuchung von Bedingungsfaktoren* (Dissertation, Universität Duisburg-Essen). Zugriff am 2018-15-08 auf https://core.ac.uk/download/pdf/33798677.pdf

Dörsam, M. & Lauber, V. (2015). *The effect of a compressed high school curriculum on university performance.* Kiel und Hamburg: ZBW-Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft.

Freyer, K., Epple, M., Brand, M., Schiebener, J. & Sumfleth, E. (2014). Studienerfolgsprognose bei Erstsemesterstudierenden in Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, *20* (1), 129–142.

Fries, M. (2002). Abitur und Studienerfolg. Welchen „Wert“ hat das Abitur für ein erfolgreiches Studium? *Beiträge zur Hochschulforschung*, *24* (1), 30–51.

Georg-August-Universität Göttingen. (2018). *Studierende Verlaufsdaten.* Zugriff am 2018-08-08 auf http://www.uni-goettingen.de/de/24653.html

Gignac, G. E. & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and individual differences*, *102*, 74–78.

Gold, A. & Souvignier, E. (2005). Prognose der Studierfähigkeit. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, *37* (4), 214–222.

Gueudet, G. (2008). Investigating the secondary–tertiary transition. *Educational studies in mathematics*, *67* (3), 237–254.

Halverscheid, S. & Pustelnik, K. (2013). Studying math at the university: Is dropout predictable. In *Proceedings of the 37th conference of the international group for psychology of mathematics education* (Bd. 2, S. 417–424).

Halverscheid, S., Pustelnik, K., Schneider, S. & Taake, A. (2014). Ein diagnostischer Ansatz zur Ermittlung von Wissenslücken zu Beginn mathematischer Vorkurse. In *Mathematische vor-und brückenkurse* (S. 295–308). Springer.

Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für pädagogische Psychologie*, *25* (2), 89–103.

Hartung, J., Knapp, G. & Sinha, B. K. (2011). *Statistical meta-analysis with applications* (Bd. 738). John Wiley & Sons.

Helbig, M., Jähnen, S. & Marczuk, A. (2017). Eine Frage des Wohnorts. *Zeitschrift für Soziologie*, *46* (1), 55–70.

Hell, B., Trapmann, S. & Schuler, H. (2007). Eine Metaanalyse der Validität von fachspezifischen Studierfähigkeitstests im deutschsprachigen Raum. *Empirische Pädagogik*, *21* (3), 251–270.

Heller, K. & Perleth, C. (2000). *Kognitiver fähigkeitstest für 4. bis 12. klassen, revision.[cognitive abilities test (cogat; thorndike, l. & hagen, e., 1954-1986)–german adapted version/author)].* Göttingen: Beltz. Google Scholar.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58* (1), 78–79.

Henoch, J. R., Klusmann, U., Lüdtke, O. & Trautwein, U. (2015). Who becomes a teacher? Challenging the "negative selection" hypothesis. *Learning and Instruction*, *36*, 46–56.

Hessisches Kultusministerium. (2011). *HKM: Leistungen von G-8 Schülern nicht schlechter als die von G9-Schülern.* Zugriff am 2018-07-11 auf http://gymnasium.bildung.hessen.de/news/news_item_1307077849.html

Heublein, U., Richter, J., Schmelzer, R. & Sommer, D. (2012). Die Entwicklung der Schwund-und Studienabbruchquoten an den deutschen Hochschulen. In *HIS: Forum Hochschule* (Bde. 3, p. 2012).

Hiebert, J. & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. *Conceptual and procedural knowledge: The case of mathematics*, *2*, 1–27.

Hoe, N. D. (2014). *Not all types of delay are equal: Postsecondary delay in the US and taking a gap year* (Dissertation, University of Pennsylvania). Zugriff am 2018-15-08 auf https://repository.upenn.edu/cgi/viewcontent.cgi?article=3125&context=edissertations

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments.* Psychological Assessment Resources.

Holmeier, M. (2012). Vergleichbarkeit der Punktzahlen im schriftlichen Abitur. In *Zentralabitur* (S. 293–324). Springer.

Homuth, C. (2017). *Die G8-Reform(en) in Deutschland.* Springer.

Hoppenbrock, A., Biehler, R., Hochmuth, R. & Rück, H.-G. (2016). *Lehren und Lernen von Mathematik in der Studieneingangsphase.* Springer.

Horn, W. (1983). Lps leistungsprüfsystem. *Göttingen: Hogrefe*.

Hübner, N., Wagner, W., Kramer, J., Nagengast, B. & Trautwein, U. (2017). Die G8-Reform in Baden-Württemberg: Kompetenzen, Wohlbefinden und Freizeitverhalten vor und nach der Reform. *Zeitschrift für Erziehungswissenschaft*, *20* (4), 748–771.

Iannone, P. & Nardi, E. (2007). The interplay between syntactic and semantic knowledge in proof production: Mathematicians perspectives. In *Proceedings of the fifth congress of the european society for research in mathematics education* (S. 2300–2309).

Kahnert, J. (2014). *Das Zentralabitur im Fach Mathematik: Eine empirische Analyse von Abitur-und TIMSS-Daten im Vergleich* (Bd. 47). Waxmann Verlag.

Kaub, K., Karbach, J., Biermann, A., Friedrich, A., Bedersdorfer, H.-W., Spinath, F. M. & Brünken, R. (2012). Berufliche Interessensorientierungen und kognitive Leistungsprofile von Lehramtsstudierenden mit unterschiedlichen Fachkombinationen. *Zeitschrift für pädagogische Psychologie*, *26* (4), 233–249.

Kersten, I. (2005). *Analytische Geometrie und Lineare Algebra* (Bd. 1). Universitätsverlag Göttingen.

Klusmann, U., Trautwein, U., Lüdtke, O., Kunter, M. & Baumert, J. (2009). Eingangsvoraussetzungen beim Studienbeginn: Werden die Lehramtskandidaten unterschätzt? *Zeitschrift für Pädagogische Psychologie*, *23* (34), 265–278.

Kühn, S. M. (2014). Sind 12 Schuljahre ausreichend für den Zugang zur Hochschule? Der doppelte Abiturjahrgang aus empirischer Perspektive. *Beiträge zur Hochschulforschung*, *36* (3), 8–33.

Kultusministerkonferenz. (2002). *Einheitliche prüfungsanforderungen im fach mathematik.* Zugriff am 2018-15-08 auf https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Mathe.pdf

Kultusministerkonferenz. (2004). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (Jahrgangsstufe 10).* Zugriff am 2018-15-08 auf https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-Mathe-Mittleren-SA.pdf

Kultusministerkonferenz. (2012). *Bildungsstandards im Fach Mathematik für die allgemeine Hochschulreife.* Educational standards in the subject mathematics for the general qualification for university entrance), Berlin: Kultusministerkonferenz. Zugriff am 2018-15-08 auf https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Mathe-Abi.pdf

Leucht, M., Kampa, N. & Köller, O. (2016). *Fachleistungen beim Abitur: Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein.* Waxmann Verlag.

Levine, T. R. & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, *28* (4), 612–625.

Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational studies in mathematics*, *52* (1), 29–55.

Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, *54* (1), 17–24.

Meyer, T. & Thomsen, S. L. (2016). *How important is secondary school duration for post-school education decisions? Evidence from a natural experiment.* Diskussionspapiere der Wirtschaftswissenschaftlichen Fakultät, Universität Hannover, No. 509, Wirtschaftswissenschaftliche Fakultät, Universität Hannover, Hannover.

Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S. & Poskowsky, J. (2017). *Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. 21. Sozialerhebung des Deutschen Studentenwerks – durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung. Berlin: Bundesministerium für Bildung und Forschung (BMBF).*

Middendorff, E., Apolinarski, B., Poskowsky, J., Kandulla, M. & Netz, N. (2013). Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2012.20. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch das HIS-Institut für Hochschulforschung. *Sozialerhebung des Deutschen Studentenwerks durchgeführt durch das HIS-Institut für Hochschulforschung. Berlin.*

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2013). *Ministerin Löhrmann: NRW hat Abitur mit Doppeljahrgang erfolgreich bewältigt.* Zugriff am 2018-07-11 auf https://bildungsklick.de/bundeslaender/meldung/ministerin-loehrmann-nrw-hat-abitur-mit-doppeljahrgang-erfolgreich-bewaeltigt/

Nagy, G. (2007). *Berufliche Interessen, kognitive und fachgebundene Kompetenzen* (Dissertation, Freie Universität Berlin). Zugriff am 2018-15-08 auf https://refubium.fu-berlin.de/handle/fub188/10012

Nardi, E. (2007). *Amongst mathematicians: Teaching and learning mathematics at university level* (Bd. 3). Springer Science & Business Media.

Nardi, E. (2011). Driving noticing'yet 'risking precision': University mathematicians' pedagogical perspectives on verbalisation in mathematics. In *Proceedings of the 7th conference of european researchers in mathematics education* (S. 2053–2062).

Neumann, M., Nagy, G., Trautwein, U. & Lüdtke, O. (2009). Vergleichbarkeit von Abiturleistungen. *Zeitschrift für Erziehungswissenschaft*, *12* (4), 691–714.

Niedersächsisches Kultusministerium. (2006). *Kernkurriculum für das Gymnasium 10 Schuljahrgänge 5-10, Mathematik, Niedersachsen.* Zugriff am 2018-15-08 auf http://db2.nibis.de/1db/cuvo/datei/kc_gym_mathe_nib.pdf

Niedersächsisches Kultusministerium. (2009). *Kerncurriculum für das Gymnasium- gymnasiale Oberstufe, Mathematik.* Hannover. Zugriff am 2018-15-08 auf http://db2.nibis.de/1db/cuvo/datei/kc_mathematik_go_i_2009.pdf

Niedersächsisches Kultusministerium. (2011). *Auswertung Zentralabitur 2011 - Notenspiegel.* Zugriff am 2018-07-11 auf http://www.nibis.de/nli1/allgemein/gosin/zentralabitur/auswertung2011/Abitur_Auswertung_2011_Notenspiegel_Nds_Region.pdf

Niedersächsisches Kultusministerium. (2016a). *Auswertung Zentralabitur 2016 - Notenspiegel.* Zugriff am 2018-07-11 auf http://www.nibis.de/nli1/allgemein/gosin/zentralabitur/auswertung2016/Abitur_Auswertung_2016_Notenspiegel_%20Nds_Region.pdf

Niedersächsisches Kultusministerium. (2016b). *Auswertung Zentralabitur 2016 - Übersicht Fächer.* Zugriff am 2018-07-11 auf http://www.nibis.de/nli1/allgemein/gosin/zentralabitur/auswertung2016/20160721Abitur-Auswertung_2016_4_Faecher%20-%20Nds_LSchB.pdf

Niedersächsisches Kultusministerium. (2016c). *Durchschnitt 2,58: Landesweite Abiturergebnisse in Niedersachsen erneut verbessert - Ergebnisse im Fach Mathematik im Bereich der Vorjahre.* Zugriff am 2018-07-11 auf https://www.mk.niedersachsen.de/startseite/aktuelles/presseinformationen/-durchschnitt-258-landesweite-abiturergebnisse-in-niedersachsen-erneut-verbessert---ergebnisse-im-fach-mathematik-im-bereich-der-vorjahre-144986.html

Niu, S. & Tienda, M. (2013). Delayed enrollment and college plans: Is there a postponement penalty? *The Journal of higher education*, *84* (1), 1–27.

Oehrtman, M. (2009). Collapsing dimensions, physical limitation, and other student metaphors for limit concepts. *Journal for Research in Mathematics Education*, 396–426.

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (2013). IQB-Ländervergleich 2012: mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I. *Münster/New York/München/Berlin: Waxmann.*

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., … Schiefele, U. (2005). *PISA 2003: der zweite Vergleich der Länder in Deutschland-was wissen und können Jugendliche?* Waxmann.

Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012: Fortschritte und Herausforderungen in Deutschland.* Waxmann Verlag.

Rach, S. & Heinze, A. (2013). Welche Studierenden sind im ersten Semester erfolgreich? *Journal für Mathematik-Didaktik*, *34* (1), 121–147.

Rach, S. & Heinze, A. (2017). The transition from school to university in mathematics: Which influence do school-related variables have? *International Journal of Science and Mathematics Education*, *15* (7), 1343–1363.

Rasch, G. (1960). *Studies in mathematical psychology: I. probabilistic models for some intelligence and attainment tests.* Nielsen & Lydiche.

Roksa, J. & Velez, M. (2012). A late start: Delayed entry, life course transitions and bachelor's degree completion. *Social forces*, *90* (3), 769–794.

Rost, J. (1996). *Lehrbuch testtheorie, testkonstruktion.* H. Huber.

Ruxton, G. D. & Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behavioral Ecology*, *19* (3), 690–693.

Schmider, E., Ziegler, M., Danay, E., Beyer, L. & Bühner, M. (2010). Is it really robust? *Methodology*.

Schneider, H., Franke, B., Woisch, A. & Spangenberg, H. (2017). *Erwerb der Hochschulreife und nachschulische Übergänge von Studienberechtigten: Studienberechtigte 2015 ein halbes Jahr vor und ein halbes Jahr nach Schulabschluss*. Deutsches Zentrum für Hochschul-und Wissenschaftsforschung (DZHW).

Sibbertsen, P. & Stöver, B. (2017). *Die räumliche Flexibilität von Studierenden: Gründe für das Wanderungsverhalten von Studienanfänger/-innen zwischen den Bundesländern.* Hannover Economic Papers (HEP), No. 604, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät, Hannover.

Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Hrsg.), *On the teaching of linear algebra* (S. 209–246). Springer.

Sorge, S., Petersen, S. & Neumann, K. (2016). Die Bedeutung der Studierfähigkeit für den Studienerfolg im 1. Semester in Physik. *Zeitschrift für Didaktik der Naturwissenschaften*, *22* (1), 165–180.

Stadler, E. (2011). The same but different—novice university students solve a textbook exercise. In M. Pytlak, T. Rowland & E. Swoboda (Hrsg.), *Proceedings of the 7th conference of european researchers in mathematics education* (S. 2083–2092).

Statista. (o. J.). *Durchschnittliche Abiturnoten in Deutschland nach Bundesländern im Schuljahr 2015/2016.* Zugriff am 2018-07-11 auf https://de.statista.com/statistik/daten/studie/36277/umfrage/durchschnittliche-abiturnoten-im-vergleich-der-bundeslaender/

Stewart, S. & Thomas, M. O. (2009). A framework for mathematical thinking: The case of linear algebra. *International Journal of Mathematical Education in Science and Technology*, *40* (7), 951–961.

Stewart, S. & Thomas, M. O. (2010). Student learning of basis, span and linear independence in linear algebra. *International Journal of Mathematical Education in Science and Technology*, *41* (2), 173–188.

Strobl, C. (2015). *Das rasch-modell: Eine verständliche einführung für studium und praxis.* Rainer Hampp Verlag.

Tall, D. (1991). *Advanced mathematical thinking* (Bd. 11). Springer Science & Business Media.

Thomas, M. O., de Freitas Druck, I., Huillet, D., Ju, M.-K., Nardi, E., Rasmussen, C. & Xie, J. (2015). Key mathematical concepts in the transition from secondary school to university. In *The proceedings of the 12th international congress on mathematical education* (S. 265–284).

Trapmann, S., Hell, B., Hirn, J.-O. W. & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift für Psychologie/Journal of Psychology*, *215* (2), 132–151.

Trapmann, S., Hell, B., Weigand, S. & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs-eine Metaanalyse. *Zeitschrift für pädagogische Psychologie*, *21* (1), 11–27.

Vandebrouck, F. (2011). Students' conceptions of functions at the transition between secondary school and university. In *Proceedings of the 7th conference of european researchers in mathematics education* (S. 2093–2102).

Van Gennep, A. (2013). *The rites of passage.* Routledge.

Vieluf, U., Ivanov, S. & Nikolova, R. (2011). *Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe.* Münster: Waxmann.

Warwas, J. (2008). *Leistungsentwicklungen und berufliche Interessen in der gymnasialen Oberstufe* (Dissertation, Niedersächsische Staats-und Universitätsbibliothek Göttingen). Zugriff am 2018-15-08 auf http://ediss.uni-goettingen.de/bitstream/handle/11858/00-1735-0000-0006-AD18-7/warwas.pdf?sequence=1

Welch, B. L. (1947). The generalization of student's problem when several different population variances are involved. *Biometrika*, *34* (1/2), 28–35.

Witt, I. (2012). *Differential- und Integralrechnung 1.* Zugriff am 2018-07-11 auf http://www.uni-math.gwdg.de/iwitt/Skript.pdf

# 12 Anhang

## 12.1 Testmodellierung

| | $LR$ | $df$ | $p$ |
|---|---|---|---|
| Alg | 12,922 | 19 | 0,843 |
| Glei | 12,847 | 12 | 0,380 |
| Poly | 14,924 | 8 | 0,061 |
| Exp | 5,867 | 13 | 0,951 |
| Trig | 9,208 | 13 | 0,757 |
| Vec | 15,840 | 15 | 0,393 |
| Diff | 17,695 | 16 | 0,342 |
| Int | 14,970 | 14 | 0,380 |

Tabelle 31: Andersen-LR-Test für die globale Überprüfung der Personenhomogenität

| | $ML$ | $df$ | $p$ |
|---|---|---|---|
| Alg | 78,04 | 99 | 0,94 |
| Glei | 39,61 | 41 | 0,53 |
| Poly | 16,00 | 19 | 0,66 |
| Exp | 37,96 | 48 | 0,85 |
| Trig | 37,96 | 48 | 0,85 |
| Vec | 48,04 | 63 | 0,92 |
| Diff | 141,07 | 63 | 0,99 |
| Int | 68,59 | 56 | 0,10 |

Tabelle 32: Martin-Löf-Test für die globale Überprüfung der Itemhomogenität

| | 1-dim Rasch-Modell | Birnbaum-Modell | mixed Rasch-Modell (2 Gruppen) | mixed Rasch-Modell (3 Gruppen) |
|---|---|---|---|---|
| Alg | 5486,20 | 5574,89 | 5625,878 | 5777,662 |
| Glei | 3589,76 | 3610,46 | 3637,960 | 3717,847 |
| Poly | 2232,65 | 2232,39 | 2097,753 | 2181,766 |
| Exp | 4178,05 | 4241,20 | 4120,830 | 4207,803 |
| Trig | 3770,47 | 3812,97 | 3784,396 | 3896,148 |
| Vec | 4072,88 | 4107,85 | 4077,070 | 4121,452 |
| Diff | 4534,45 | 4613,41 | 4502,203 | 4634,658 |
| Int | 4096,39 | 4118,30 | 4098,781 | 4203,086 |

Tabelle 33: Informationskriterium BIC für den Vergleich verschiedener Modelle

## 12.2 Effektstärken der Kursform im Vergleich für die Testbereiche

| Test | erhöhtes Anforderungdniveaus | grundlegendes Anforderungdniveaus | Effekstärke |
|---|---|---|---|
| Grundlagen der Algebra | 1,90 | 1,35 | $d = 0,43$ |
| Gleichungen und Ungleichungen | 1,44 | 1,07 | $d = 0,35$ |
| Polynome | 1,80 | 1,20 | $d = 0,33$ |
| Exponential- und Logarithmus-funktionen | 0,96 | 0,64 | $d = 0,33$ |
| Trigonometrie | 0,82 | 0,01 | $d = 0,48$ |
| Vektor-rechnung | 1,58 | 0,73 | $d = 0,55$ |
| Differential-rechnung | 1,42 | 0,76 | $d = 0,56$ |
| Integral-rechnung | 1,14 | 0,66 | $d = 0,39$ |

Tabelle 34: Mittelwerte und Effekstärken des Kursform im Gruppenvergleich für die Testbereiche; alle Mittelwertunterschiede sind signifikant mit $p \leq 0,001$

## 12.3 Bestehen der Klausuren unter Verwendung aller Variablen

| Prädiktor | Differential- und Integralrechnung 1 $p$ | Analytische Geometrie und lineare Algebra 1 $p$ |
|---|---|---|
| Konstante | 0,147 | 0,938 |
| Testergebnis | 0,007 | 0,130 |
| Abiturnote | 0,115 | 0,223 |
| letzte Mathenote | 0,108 | 0,295 |
| Studiengang_Ph | 0,151 | 0,278 |
| Studiengang_LA | 0,289 | 0,441 |
| Geschlecht_weiblich | 0,422 | 0,597 |
| Schuldauer_13Jahre | 0,016 | 0,984 |
| Verzögerung_1 | 0,564 | 0,909 |
| Verzögerung_$\geq 2$ | 0,942 | 0,564 |
| Bundesland_Hessen | 0,089 | 0,686 |
| Bundesland_NRW | 0,080 | 0,184 |
| Bundesland_Schl-Hols | 0,013 | 0,999 |
| Kursform_LK | 0,682 | 0,934 |
| Jahrgang | 0,011 | 0,172 |

Tabelle 35: Logistische Regression für das Bestehen beider Klausuren des ersten Semesters unter Einschluss alle Variablen; Regressionskoeffizienten sind aufgrund der Signifikanz nicht angegeben( $p$: Wahrscheinlichekeit für die Verschiedenheit von 0)

## 12.4 Zensuren in den Klausuren unter Verwendung aller Variablen

### 12.4.1 Differential- und Integralrechnung 1

| Prädiktor | $t$ | $p$ |
|---|---|---|
| Konstante | 3,560 | $\leq 0,001$ |
| Testergebnis | -4,065 | $\leq 0,001$ |
| Abiturnote | 1,735 | 0,086 |
| letzte Mathenote | -0,927 | 0,356 |
| Studiengang_Ph | -0,410 | 0,683 |
| Studiengang_LA | -0,161 | 0,873 |
| Geschlecht_weiblich | 1,290 | 0,200 |
| Schuldauer_13Jahre | -1,735 | 0,086 |
| Verzögerung_1 | -0,884 | 0,379 |
| Verzögerung_$\geq 2$ | 0,342 | 0,733 |
| Bundesland_Hessen | 0,928 | 0,356 |
| Bundesland_NRW | 1,051 | 0,296 |
| Bundesland_Schl-Hols | 2,324 | 0,022 |
| Kursform_LK | 0,168 | 0,962 |
| Jahrgang | -3,382 | 0,001 |

Tabelle 36: Ergebnisse der Regressionsnalyse für die Veranstaltung "Differential- und Integralrechnung 1" unter Einschluss aller Variablen in das Modell; die Regressionskoeffizienten sind aufgrund der Signifikanz nicht angegeben; das Modell besitzt ein Bestimmtheitsmaß von $R^2 = 0,555$

### 12.4.2 Analytische Geometrie und lineare Algebra 1

| Prädiktor | $t$ | $p$ |
|---|---|---|
| Konstante | 4,112 | $\leq 0,001$ |
| Testergebnis | -3,441 | $\leq 0,001$ |
| Abiturnote | 1,622 | 0,108 |
| letzte Mathenote | -1,679 | 0,096 |
| Studiengang_Ph | 2,217 | 0,029 |
| Studiengang_LA | 0,522 | 0,603 |
| Geschlecht_weiblich | 0,963 | 0,338 |
| Schuldauer_13Jahre | -0,317 | 0,752 |
| Verzögerung_1 | -1,484 | 0,141 |
| Verzögerung_$\geq 2$ | -0,039 | 0,969 |
| Bundesland_Hessen | 0,472 | 0,638 |
| Bundesland_NRW | -0,895 | 0,373 |
| Bundesland_Schl-Hols | -0,644 | 0,521 |
| Kursform_LK | 0,168 | 0,867 |
| Jahrgang | 1,101 | 0,274 |

Tabelle 37: Ergebnisse der Regressionsnalyse für die Veranstaltung "Analytische Geometrie und lineare Algebra 1" unter Einschluss aller Variablen in das Modell; die Regressionskoeffizienten sind aufgrund der Signifikanz nicht angegeben; das Modell besitzt ein Bestimmtheitsmaß von $R^2 = 0,478$

## 12.5 Bestehen der Klausuren unter Verwendung aller Gebiete

### 12.5.1 Differential und Integralrechnung 1

| Test | signifikante Prädiktoren | Mc Faddens $R^2$ | richtig zugeordnete Fälle |
|---|---|---|---|
| Grundlagen der Algebra | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,22 | 81,3 % |
| Gleichungen und Ungleichungen | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,20 | 79,9 % |
| Polynome | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,22 | 79,9 % |
| Exponential- und Logarithmus-funktionen | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,24 | 82,8 % |
| Trigonometrie | letzte Mathenote<br>Jahrgang | 0,17 | 79,1 % |
| Vektor-rechnung | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,24 | 82,1 % |
| Differential-rechnung | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,37 | 84,3% |
| Integral-rechnung | letzte Mathenote<br>Testergebnis<br>Jahrgang | 0,21 | 77,6 % |

Tabelle 38: Ergebnisse der logistischen Regressionsnalysen für die Veranstaltung "Differential- und Integralrechnung 1" unter Verwendung der Teiltestergebnisse, der Abiturnote, der letzten Mathenote und des Jahrgangs; angegeben sind die in das Regressionsmodell aufgenommenen Variablen sowie die jeweiligen Pseudo-estimmtheitsmaße und die Anzahl korrekt zugeordneter Fälle mit dem Trennwert 0,5

### 12.5.2 Analytische Geometrie und lineare Algebra 1

| Test | signifikante Prädiktoren | Mc Faddens $R^2$ | richtig zugeordnete Fälle |
|---|---|---|---|
| Grundlagen der Algebra | Abiturnote | 0,14 | 75,2 % |
| Gleichungen und Ungleichungen | Abiturnote<br>Testergebnis | 0,18 | 79,4 % |
| Polynome | Abiturnote<br>Testergebnis | 0,20 | 78,0 % |
| Exponential- und Logarithmus-funktionen | Abiturnote<br>Testergebnis | 0,18 | 78,8 % |
| Trigonometrie | Abiturnote | 0,14 | 75,2 % |
| Vektor-rechnung | Abiturnote | 0,14 | 75,2 % |
| Differential-rechnung | Abiturnote<br>Testergebnis | 0,19 | 76,6 % |
| Integral-rechnung | Abiturnote | 0,14 | 75,2 % |

Tabelle 39: Ergebnisse der logistischen Regressionsnalysen für die Veranstaltung "Analytische Geometrie und lineare Algebra 1" unter Verwendung der Teiltestergebnisse, der Abiturnote, der letzten Mathenote und des Jahrgangs; angegeben sind die in das Regressionsmodell aufgenommenen Variablen sowie die jeweiligen Pseudo-Bestimmtheitsmaße und die Anzahl korrekt zugeordneter Fälle mit dem Trennwert 0,5

## 12.6 Zensuren der Klausuren mit allen Gebieten

### 12.6.1 Differential und Integralrechnung 1

| Test | signifikante Prädiktoren | Bestimmtheitsmaß $R^2$ |
|---|---|---|
| Grundlagen der Algebra | Abiturnote<br>Testergebnis | 0,362 |
| Gleichungen und Ungleichungen | Abiturnote<br>Jahrgang<br>Testergebnis | 0,376 |
| Polynome | Abiturnote<br>Jahrgang<br>Testergebnis | 0,385 |
| Exponential- und Logarithmus-funktionen | Testergebnis<br>letzte Mathenote<br>Jahrgang<br>Abiturnote | 0,400 |
| Trigonometrie | Abiturnote<br>Jahrgang<br>Testergebnis | 0,396 |
| Vektor-rechnung | Testergebnis<br>letzte Mathenote<br>Jahrgang<br>Abiturnote | 0,436 |
| Differential-rechnung | Testergebnis<br>Jahrgang<br>Abiturnote | 0,448 |
| Integral-rechnung | Abiturnote<br>Jahrgang<br>Testergebnis | 0,368 |

Tabelle 40: Ergebnisse der linearen Regressionsnalysen für die Veranstaltung "Differential- und Integralrechnung 1" unter Verwendung der Teiltestergebnisse, der Abiturnote, der letzten Mathenote und des Jahrgangs; angegeben sind die in das Regressionsmodell aufgenommenen Variablen sowie die jeweiligen Bestimmtheitsmaße

### 12.6.2 Analytische Geometrie und lineare Algebra 1

| Test | signifikante Prädiktoren | Bestimmtheitsmaß $R^2$ |
|---|---|---|
| Grundlagen der Algebra | Abiturnote<br>letzte Mathenote | 0,371 |
| Gleichungen und Ungleichungen | Abiturnote<br>Testergebnis<br>letzte Mathenote | 0,439 |
| Polynome | Abiturnote<br>letzte Mathenote<br>Testergebnis | 0,405 |
| Exponential- und Logarithmus-funktionen | Abiturnote<br>letzte Mathenote<br>Testergebnis | 0,398 |
| Trigonometrie | Abiturnote<br>letzte Mathenote<br>Testergebnis | 0,399 |
| Vektor-rechnung | Abiturnote<br>letzte Mathenote<br>Testergebnis | 0,398 |
| Differential-rechnung | Abiturnote<br>Testergebnis<br>letzte Mathenote | 0,412 |
| Integral-rechnung | Abiturnote<br>letzte Mathenote | 0,371 |

Tabelle 41: Ergebnisse der linearen Regressionsnalysen für die Veranstaltung "Analytische Geometrie und lineare Algebra 1" unter Verwendung der Teiltestergebnisse, der Abiturnote, der letzten Mathenote und des Jahrgangs; angegeben sind die in das Regressionsmodell aufgenommenen Variablen sowie die jeweiligen Bestimmtheitsmaße

**Curriculum Vitae**

Kolja Pustelnik
07.10.1988 in Leer, Niedersachsen

| | |
|---|---|
| September 2014 – August 2015 | Lehrkraft an der Berufsbildenen Schule Ritterplan |
| seit Mai 2013 | Promotionsstudent im Promotionsstudiengang Mathematical Sciences |
| seit Oktober 2013 | Wissenschaftlicher Mitarbeiter am Mathematischen Institut, Georg-August-Universität Göttingen |
| September 2013<br>Masterarbeit | Master of Education<br>Ermittlung von Wissenslücken zu Beginn des mathematischen Studiums: ein diagnostischer Ansatz zur pädagogischen Förderung |
| September 2011 - September 2013 | Studium im Master of Education mit den Fächern Mathematik und Physik |
| August 2010<br>Bachelorarbeit | Bachelor of Arts<br>Der Satz von Gauß über faktorielle Ringe |
| Oktober 2007 - August 2010 | Studium 2-Fächer Bachelor mit den Fächern Mathematik und Physik im Profil Lehramt |
| Juni 2007 | Allgemeine Hochschulreife, Ubbo-Emmius-Gymnasium Leer |

**Publikation im Zusammenhang mit der Dissertation:**

Halverscheid, S. & Pustelnik, K. (2013). Studying math at the university: Is dropout predictable. In Proceedings of the 37th Conference of the International Group for Psychology of Mathematics Education (Vol. 2, pp. 417-424).

Halverscheid, S., Pustelnik, K., Schneider, S. & Taake, A. (2014). Ein diagnostischer Ansatz zur Ermittlung von Wissenslücken zu Beginn mathematischer Vorkurse. In Mathematische Vor-und Brückenkurse (pp. 295-308). Springer Spektrum, Wiesbaden.

Pustelnik, K. (2014). Variables influencing the competencies of first-year students at the university. In Liljedahl, P., Nicol, C., Oesterle, S. & Allan, D. (Eds.). Proceedings of the Joint Meeting of PME 38 and PME-NA 36 (Vol. 6). Vancouver, Canada (p.201): PME.

Pustelnik, K. Halverscheid, S. (2016). On the consolidation of declarative mathematical knowledge at the transition to tertiary education. In Csíkos, C., Rausch, A. & Szitányi, J. (Eds.). Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education, Vol. 2. Szeged, Hungary: PME.