

Vol. 3, No. 2
Oktober 2012

Anna Wolf, Reinhard Kopiez & Friedrich Platz

Hochschule für Musik, Theater und Medien Hannover

**Der Status Quo der musiktheoretischen Zulassungsprüfung
an Musikhochschulen: Eine testtheoretische Analyse**

**The status quo of the collegiate entrance exam in music theory
analyzed from the perspective of test theory**

[Artikel als PDF](#)

[urn:nbn:de:101:1-2012110811327](https://nbn-resolving.org/urn:nbn:de:101:1-2012110811327)

© Wolf, Kopiez & Platz 2012 All rights reserved

Abstract

Die Zulassungsprüfung an Musikhochschulen ist von zentraler Bedeutung für Studieninteressierte, da sie als Entscheidungsgrundlage für die Aufnahme in einen musikpraktischen Studiengang herangezogen wird. Ihr Ziel ist der Nachweis über „eine besondere künstlerische Befähigung“ (Hochschule für Musik, Theater und Medien Hannover, o. J.). Zulassungsprüfungen beinhalten neben einem musikpraktischen auch einen musiktheoretischen

Teil, der in der Regel durch eine jährlich neu konzipierte Klausur durchgeführt wird.

In dieser Studie wird das Ergebnis einer statistischen Analyse der Musiktheorieklausuren ($n = 302$) der Jahre 2009 bis 2011 an einer deutschen Musikhochschule vorgestellt. Hierbei werden die statistischen Kennwerteverteilungen der einzelnen Aufgaben gemäß testtheoretischer Kriterien untersucht und bewertet. Unter Verwendung eines probabilistischen Testmodells (Rasch-Modell) wurde exemplarisch die psychometrische Qualität einer Akkordbenennungsaufgabe untersucht. Die Aufgabenanalyse hat gezeigt, dass die Items zwar größtenteils Rasch-konform sind, jedoch eine zu geringe Schwierigkeitsspanne abdecken. Da außerdem andere Prüfungsteilaufgaben nicht objektiv zu bewerten und somit statistisch nicht prüfbar sind, liegt nahe, dass bei der Aufgabenkonstruktion für die Zulassungsprüfung gängige Richtlinien der Testentwicklung (Hornke, 2005) unberücksichtigt bleiben.

Wir sehen unsere statistische Analyse als Ausgangspunkt für die Entwicklung kompetenzorientierter Aufgaben bei gleichzeitiger Entwicklung eines Kompetenzmodells für die Fächer Musiktheorie und Gehörbildung. Langfristig wird angeregt, einen Aufgabenpool zu entwickeln, der für die Zulassungsprüfung in Musiktheorie und Gehörbildung an Hochschulen Rasch-Modell-konforme Aufgaben zur Verfügung stellt. Hierdurch könnten fachimmanente Standards definiert werden, die es jungen Musikern ermöglichen, sich zielgerichtet auf eine Zulassungsprüfung vorzubereiten.

Schlagwörter: Eignungsdiagnostik, Item response theory, Musiktheorie

Summary

The entrance exam at music colleges is of pivotal importance as its results regulate if the prospective student may or may not enter the chosen program of studies. The objective of entrance exams is to provide evidence of "an exceptional artistic qualification" (Hanover University of Music, Drama and Media, n. d.)¹. Proficiency in both musical practice and music theory are tested, the latter in a written exam. Such a written exam is developed from scratch every year and tests distinct knowledge in music theory and aural training.

This study presents the results of a statistical analysis of music theory entrance exams from a German music college. Having accessed a comprehensive sample of entrance exam candidates (overall $n = 302$), we analyzed the distributions of the results from the perspective of test theory. The item analysis has shown that the items are mostly Rasch-conform, however, the item difficulties are within a too narrow range. Additionally, the other tasks cannot be marked objectively and are therefore not analyzable. This provides evidence that the task design does not conform with established standards in test design (Hornke, 2005).

This statistical analysis is the starting point for the prospective development of competency-oriented items and the advancement of a competency model for music theory and ear training. On a long-term basis we suggest the development of an item pool which could supply theoretical based test items for the entrance exam in music theory and aural training. Such an improvement would define subject-specific standards and enable young musicians to prepare themselves for an entrance exam in an optimized and goal-oriented way.

Keywords: Item reponse theory, Music theory, Proficiency assessment

¹ translated by Anna Wolf

1 Einleitung

Im Vorfeld der Immatrikulation in musikpraktische, musikpädagogische und manche musikwissenschaftliche Studiengänge findet eine Aufnahme-, Eignungs- bzw. Zulassungsprüfung² statt. Sie bietet die Entscheidungsgrundlage dafür, ob ein Bewerber in den gewünschten Studiengang aufgenommen wird. Zu diesem Zweck werden die künstlerisch-praktischen und musiktheoretischen Fertigkeiten des Bewerbers erfasst (Rektorenkonferenz der deutschen Musikhochschulen in der HRK [RKM], 2009).

Diese fachspezifische zusätzliche Zulassungsprüfung stellt eine Spezialform der Zugangsberechtigung an deutschen Hochschulen und Universitäten dar, die unter anderem für künstlerisch-wissenschaftliche bzw. rein künstlerische Studiengänge gilt und jeweils vor Ort in einer von jeder Hochschule frei gestalteten Form durchgeführt wird.

1.1 Anforderungen in den musiktheoretischen Zulassungsprüfungen verschiedener Hochschulen

Das Auswahlverfahren an Musikhochschulen und Universitäten wird rechtlich von studienengangsspezifischen Zulassungsprüfungsordnungen geregelt. Wer sich beispielsweise an der Hochschule für Musik, Theater und Medien (HMTM) in Hannover für das Lehramt an Gymnasien ausbilden lassen möchte, kann nur dann das Studium aufnehmen, wenn eine besondere künstlerische Befähigung nachgewiesen werden kann (Hochschule für Musik, Theater und Medien, Hannover [HMTMH], o. J.). An der Hochschule für Künste in Bremen werden die Voraussetzungen zur Aufnahme eines Studiums detaillierter definiert. Hier wird überprüft, ob „die künstlerischen Fähigkeiten, die musiktheoretischen Vorkenntnisse und Hörfähigkeiten sowie die pädagogische Eignung für das Arbeitsfeld Musiklehrerin/Musiklehrer“ beim Bewerber vorhanden sind (Hochschule für Künste, Bremen [HfKB], 2007). An der Hochschule für Musik in Detmold wird ferner der „Nachweis einer besonderen studiengangsbezogenen musikalischen Eignung“ gefordert (Hochschule für Musik, Detmold [HfMD], 2007).

Diese drei Zulassungsprüfungen für lehramtsbezogene Studiengänge wurden exemplarisch aus drei Musikhochschul-Standorten des norddeutschen Raums ausgewählt. Voraussetzung für die Auswahl war die Verfügbarkeit einer Musterzulassungsprüfung und einer Zulassungsordnung als Quelle für die inhaltlichen und formalen Vorgaben. Die Zulassungsprüfungen in Bremen und Detmold prüfen die Fertigkeiten der Bewerber durch ein Vorspiel auf

² Ab sofort wird immer von der Zulassungsprüfung gesprochen, die genauen Bezeichnungen können an den verschiedenen Studienstandorten abweichen.

deren Haupt- und Nebeninstrumenten, in Musiktheorie und Gehörbildung anhand einer Klausur sowie anhand einer musikalisch-praktischen Gruppenanleitung. Zusätzlich wird im Rahmen der Bremer Zulassungsprüfung ein kurzes Gespräch mit den Bewerbern über die ausgewählten Stücke und das Berufsbild des Lehrers geführt. Die Zulassungsprüfung in Hannover besteht aus einer Klausur in Musiktheorie und Gehörbildung, dem Vorspiel auf den Haupt- und Nebeninstrumenten sowie einer mündlichen Prüfung in Musiktheorie und Gehörbildung. In allen drei Zulassungsprüfungen werden Gesang und ein Begleitinstrument (z. B. Klavier oder Gitarre) im musikpraktischen Prüfungsteil innerhalb der Haupt- und Nebeninstrumente geprüft.

Diese verschiedenen Prüfungsteile werden einzeln benotet. Die Gesamtnote der Zulassungsprüfung wird in der Regel als gewichteter Mittelwert errechnet. Sowohl die Gesamtnote als auch die Einzelnoten müssen eine Mindestnote (Note 4,0 in Bremen und Detmold) bzw. eine Mindestpunktzahl (7 von 15 Punkten in Hannover) erreichen. Möglichkeiten zum Ausgleich von zu niedrigen Punktzahlen sind jedoch auch gegeben: In der Detmolder Zulassungsprüfung ist eine nicht weiter erläuterte Kompensationsmöglichkeit erwähnt, wobei in Bremen eine mündliche Nachprüfung in Musiktheorie absolviert werden kann, wenn das Hauptinstrument mit der Note 1,0 benotet ist. In Hannover ist eine begrenzte Kompensation möglich, indem für eine Durchschnittsnote in Musiktheorie/Gehörbildung (Klausur und mündlich) unter 7 Punkten ein gestaffelter Punkteabzug im Hauptfach erfolgt. Ist dort ein Prüfungsteil mit null Punkten bewertet worden, gilt die gesamte Zulassungsprüfung als nicht bestanden.

In diesen und den folgenden Formulierungen ist mit „Musiktheorie“ eher die pädagogische Vermittlung (de la Motte-Haber, 2005) von Musiktheorie als die Wissenschaft gemeint, also die (Inhalte der) Lehre, die Teilbereiche wie z. B. Harmonielehre oder Kontrapunkt und häufig auch Gehörbildung umfasst, da Musiktheorie und Gehörbildung oft gleichzeitig oder in Personalunion unterrichtet werden.

Die folgende Aufstellung zeigt, dass die Hochschulen zwar einheitlich die Teilbereiche Musikpraxis und Musiktheorie prüfen, in der Musiktheorie jedoch auf unterschiedliche Arten, da an jedem Standort andere Schwerpunkte gesetzt werden. Das zeigt sich insbesondere an den online zur Verfügung gestellten Probeklausuren in Musiktheorie und Gehörbildung (siehe Tab. 1).

Tab. 1: Inhalte der Muster-Probeklausuren der Musikhochschulen Bremen, Detmold und Hannover in Musiktheorie und Gehörbildung

<i>Gehörbildung</i>	<i>Bremen</i>	<i>Detmold</i>	<i>Hannover</i>
Taktart hören	–	ja	–
Melodie wiedererkennen	–	ja, anhand Notentext (multiple choice)	–
Fehlererkennung	–	ja, in Rhythmus und polyphonem Satz	–
Intervalle hören	sukzessiv und simultan	–	–
Akkorde hören	ja, Ordnen der zuvor zu bezeichnenden Akkorde (siehe Bremen/Akkorde bezeichnen)	ja	ja
Melodiediktat	zweistimmiges melodisch-	ja	einstimmiges melodisch-rhythmisches Diktat
Rhythmusdiktat	rhythmisches Diktat	ja	
		ja, zwei-stimmiger Satz	
Höranalyse	–	ja, Bestimmung von Gattung und Epoche, geschlossenes Antwortformat	ja, drei Stücke, offenes Antwortformat
<i>Musiktheorie</i>	<i>Bremen</i>	<i>Detmold</i>	<i>Hannover</i>
Töne notieren	–	ja	–
Intervalle notieren	–	ja	–
Akkorde auflösen	ja	–	–
Akkorde bezeichnen	ja	ja	ja
Tonart bestimmen	ja	ja	–
Tonleiter/Skala notieren	–	ja	–
Melodie transponieren	–	ja	–
Melodie fortführen	–	ja	–
Satzaufgabe	–	ja, Ober- oder Unterstimme zur Melodie	ja, Harmonisierung einer Bassstimme
Partituranalyse	ja, anhand von neun Unteraufgaben, halboffenes Antwortformat	–	–
Obertonreihe notieren	–	ja	–

Quellen: HfKB, 2007; HfMD, o. J., a; HfMD, o. J., b; HMTM, o. J.

Die einzigen an allen drei Hochschulen geprüften Aufgabengebiete befassen sich mit dem Hören bzw. der Bezeichnung von Akkorden, dem Musikdiktat und der Analyse (hörend bzw. am Notentext). Diese letzte Aufgabengruppe ist jedoch im Antwortformat so unterschiedlich, dass die Aufgaben letztlich nicht vergleichbar sind. Auch in anderen Aspekten sind die Prüfungen unterschiedlich konzipiert: So werden in Bremen Aufgaben in Musiktheorie und Gehörbildung in der Klausur gemischt, in Hannover und Detmold dagegen getrennt. Die jeweiligen Prüfungen enthalten unterschiedlich viele Aufgaben und beginnen auf unterschiedlichen Schwierigkeitsebenen.

Diese große inhaltliche Varianz in der Vermittlung von Musiktheorie und Gehörbildung wird auch in einer jüngeren Untersuchung zu den Konzeptionen der Fächer, die sich in der Gehörbildungsliteratur widerspiegeln, von Estrada Rodriguez (2008) bestätigt. 13 der wichtigsten deutschen Gehörbildungslehrbücher der letzten 100 Jahre wurden inhaltlich bezüglich ihrer Aktionsformen und Aktionsformenfelder³ analysiert, wobei Estrada Rodriguez festgestellt hat, dass „bis heute in der Gehörbildungspraxis höchst individualisiert gearbeitet wird und dass diese Praxis von einer curricular standardisierten und auf Effektivität ausgerichteten Vorgehensweise noch weit entfernt ist“ (Estrada Rodriguez, 2008, S. 114f). Er erwähnt auch, dass die Gehörbildungs-Fertigkeiten des Schülers maßgeblich von der gelernten Methode abhängen und diese schließlich, abhängig von den favorisierten Methoden des Prüfers, zu unterschiedlichen Bewertungen führen. Eine Folge hieraus ist fehlende Transparenz bezüglich der Anforderungen für die Aufnahmeprüfungskandidaten, wodurch keine zielgerichtete Vorbereitung ermöglicht wird. Estrada Rodriguez' Vorschlag, diesem Problem auszuweichen, ist eine standardisierte Gehörbildungsprüfung, die hochschulübergreifend eingesetzt würde, um eine objektive Messung der Fertigkeiten zu gewährleisten.

Ein weiterer Kritikpunkt ist der festgestellte Mangel an einem inhaltlichen und methodischen Austausch: „Allgemeine didaktische Erkenntnisse, Erkenntnisse der Musikpsychologie und der Lerntheorie sind bei den untersuchten Autoren in der Reflexion nicht systematisch berücksichtigt worden“ (Estrada Rodriguez, 2008, S. 117). Die analysierten Bücher haben sich, wie z. B. auch Hempel (2010) ausführt, zum Ziel gesetzt, das Selbststudium anzuleiten oder zu unterstützen und sind somit eine wesentliche Quelle beim Erwerb musiktheoretischer Fertigkeiten⁴, insbesondere auch vor Beginn des Studiums.

³ Estrada Rodriguez definiert eine Aktionsform als „Aktivität, Aufgabe oder Übung, die aus verschiedenen Elementen [...] besteht.“ (Estrada Rodriguez, 2008, S. 13). Das Singen ist beispielsweise eine Aktionsform, aber genauso das Hören und Schreiben, was vereint die Aktionsform Diktat ausmacht. Aktionsformenfelder sind – synonym zu Inhaltsfeldern – „die einzelnen Teile der Inhaltsgliederung der Bücher“ (s. o.). Estrada Rodriguez hat insgesamt 10 Aktionsformenfelder definiert, u. a. Rhythmik, Intervalle, Akkorde, Harmonik und Formenlehre.

⁴ Von „Fertigkeiten“ wird gesprochen, wenn das jeweilige Verhalten bereits erlernt wurde und auf neue Fragestellungen anwendbar ist. Es soll also nicht diskutiert werden, ob eine Person die grundsätzliche Fähigkeit hat, Musiktheorie zu lernen, sondern welche konkreten Fertigkeiten bereits erworben wurden.

1.2 Vorbereitungsformen auf musiktheoretische Zulassungsprüfungen

Eine ähnliche Vielfalt wie beim Inhalt verschiedener Klausuren und Lehrbücher zeigt sich auch in den möglichen Formen der Vorbereitung auf eine Zulassungsprüfung, die nicht selten auf den eben erwähnten Lehrbüchern aufbaut. Die Vorbereitung auf Zulassungsklausuren findet in Ansätzen in der Kursstufe der Sekundarstufe II, in der studienvorbereitenden Ausbildung (SVA) an Musikschulen oder auch privat organisiert statt, d. h. mit Privatlehrern oder im Eigenstudium. Die folgende Darstellung beschränkt sich auf die Vorbereitung in der SVA, bei der explizit als Ziel angegeben wird, dass „der Schüler die Voraussetzungen für eine Aufnahmeprüfung an Ausbildungsstätten für Musikberufe [...] erworben“ hat (Verband deutscher Musikschulen [VdM], 1987⁵). Eine an anderer Stelle (z. B. in einem Lehr- oder Übungsbuch) veröffentlichte ähnliche „Garantie“ zur Studienaufnahme, wie sie hier formuliert wird, ist den Autoren nicht bekannt und scheint somit eine bemerkenswerte Ausnahme zu sein.

Im Jahr 2011 haben deutschlandweit 13.420 Schüler an einem Angebot aus dem Ergänzungsbereich „Musiklehre/Hörerziehung“ teilgenommen (VdM, 2011). Neben der Teilnahme an diesen Kursen besteht das Curriculum der SVA aus dem wöchentlichen Unterricht auf dem Haupt- und einem Nebeninstrument und dem Mitwirken in einem Ensemble. Außerdem hat der interessierte Schüler bereits vor Beginn der SVA eine Eignungsprüfung zu bestehen (VdM, 1987).

Neben den hauptsächlich musikpraktischen Fertigkeiten, die im Instrumentalunterricht und dem Ensemblespiel unterrichtet werden, vermittelt der Kurs Musiklehre/Hörerziehung „die Befähigung zum bewussten Umgang mit der Musik als wesentliche Hilfe für die Aufnahme und die Interpretation von Musik“ (VdM, 1990). Der Lehrplan ist in fünf je zweijährige Kurse unterteilt, wobei die Kurse 3 bis 5 obligatorisch für die SVA sind (Tab. 2).

⁵ Dieses Dokument sowie das folgende von 1990 sind trotz des frühen Zeitpunkts der Veröffentlichung immer noch die formale Grundlage für den Aufbau der SVA.

Tab. 2: Strukturplan in Hörerziehung, Musiklehre und Musiktheorie des Verbandes deutscher Musikschulen (VdM, 1990)

Stufe	Kurs-Nr.	Ausbildungszeit	Bemerkung
U I	a	1 a	1. Unterrichtsjahr
	b	1 b	2. Unterrichtsjahr
U II	2	1. Unterrichtsjahr	begrenzt obligatorisch (ggf. bis Beginn Ensemblespiel)
		2. Unterrichtsjahr	
M I	3	1. Unterrichtsjahr	fakultativ
		2. Unterrichtsjahr	obligatorisch für SVA
M II	4	1. Unterrichtsjahr	
		2. Unterrichtsjahr	
O	5	1. Unterrichtsjahr	
		2. Unterrichtsjahr	

Die Kurse 1 bis 4 sind in die Teilbereiche Gehörbildung, Musiklehre, Höranalyse, Formkunde, Improvisation und Instrumentenkunde unterteilt, der 5. Kurs dient der Wiederholung der Kurse 2 bis 4. Diese genannten Teilbereiche werden durch Listen von zu lernenden Inhalten genauer aufgeführt. Für die Teilnehmer der SVA „sollte der Aspekt des systematischen Trainings von Fähigkeiten, vor allem im Bereich der Gehörbildung, im Vordergrund stehen“ (VdM, 1990, S. 22). Zu beachten ist jedoch, dass dieser Lehrplan über 20 Jahre alt ist und es in diesen Bereichen seitdem Veränderungen gegeben haben kann, die noch nicht formalisiert wurden.

Analog zu der von Estrada Rodriguez analysierten Gehörbildungs-Literatur lässt sich auch für das SVA-Curriculum feststellen, dass durch die stagnierte Fortentwicklung auch hier keine (standardisierten) Fertigkeiten, die sich auf einem aktuellen Stand der musikpädagogischen Forschung befinden, gelehrt werden.

1.3 Musiktheoretische Fertigkeiten und ihre Messbarkeit

Innerhalb des Fachdiskurses der Musiktheorie wurde kürzlich eine Befragung an deutschen Musikhochschulen (Kühn, 2010) durchgeführt, in der eine starke Individualisierung und eine hohe Methodenvielfalt an den untersuchten Standorten festgestellt wurde, was mit den bisherigen Befunden zur (übermäßigen) Methoden- und Inhaltvielfalt in Zulassungsklausuren und der Gehörbildungsliteratur übereinstimmt. Streitpunkte haben sich u. a. bezüglich einer Kanonisierung des zu unterrichtenden Repertoires und der Wichtigkeit des Bezuges zur Mu-

sikgeschichte herauskristallisiert. Die Ergebnisse zur mangelnden Geradlinigkeit der Gehörbildungs-Didaktik von Estrada Rodriguez (2008) lassen sich somit grundsätzlich von der Lehrsituation der Gehörbildung auch auf die Lehre der Musiktheorie übertragen, da hier nur in den wenigsten Bereichen Konsens über die Inhalte und Methoden herrscht. Zuletzt wird festgestellt, dass sich das häufig ungeliebte Fach Musiktheorie die Frage stellen muss, weshalb es häufig so geringgeschätzt wird, „ob das an den störrischen Studenten liegt, die partout nicht einsehen mögen, wie bedeutsam unser Fach ist, oder ob es an dem Auftreten des Faches liegt“ (Kühn, 2010, S. 58). Auch falls Ersteres zutrifft, liegt die Verantwortung teils auf Seiten der Fachvertreter, da in jeder Lehr-Lern-Situation der Lehrende Verantwortung trägt, ob eine Lehreinheit erfolgreich abgeschlossen wird. Im Bereich der Erwachsenenbildung, zu der ein Hochschulstudium zählt, meinen die Studenten aufgrund ihrer bisherigen Erfahrung oft sehr genau zu wissen, dass Fertigkeiten in Musiktheorie und Gehörbildung nicht notwendig seien, da deren Grundkenntnisse nur selten in der vorakademischen Ausbildung vermittelt wurden (Holtmeier, 1997). Außerdem zeigt sich die erfolgreiche musikalische Weiterentwicklung auf diesem Gebiet erst langfristig (Estrada Rodriguez, 2008), was für Musikstudenten sehr demotivierend sein kann, da sie sich oft sehr leistungsorientierte Zielsetzungen stecken und eine hohe Zieltransparenz anstreben („Dieses Stück spiele ich in 3 Monaten beim Konzert!“). Für beide von Kühn erwägten Szenarien wären zunächst die zugrundeliegenden Faktoren festzustellen, die die Musiktheorie zu einem eher geringgeschätzten Fach machen, und in einem zweiten Schritt, wie dem in Zukunft zu begegnen ist.

Menke (2010) knüpft an diese Diskussion über die Lehre von Musiktheorie an und fragt, ob ein Kanon von musiktheoretischen Kompetenzen notwendig sei. Er nennt drei neu zu konzipierende Disziplinen der musiktheoretischen Lehre, die das Schreiben, Lesen, Hören und (elementare) Spielen miteinbeziehen: Kontrapunkt, Generalbass und modellorientierte Improvisation, mit denen einhergehend auch ein neuer Kanon von musikalischen Werken gebildet werden müsste. Zweck dieser Umgestaltung wäre eine verbesserte „Teilhabe an der Kultur der Tonalität“ (Menke, 2010, S. 69), wodurch die Musikstudierenden einen stärkeren Bezug zur Disziplin auch im Umgang mit ihrem Instrumentalspiel entwickeln können.

In welchem Rahmen ein Katalog von (kanonisierten) Fertigkeiten hergeleitet werden kann und wie sich dieser in der studienvorbereitenden Phase und den Studiencurricula zu unterscheiden hat, ist noch zu diskutieren. Da eine solche Diskussion mit einer Umorientierung der Messung von Fertigkeiten einhergeht, stellt sich die Frage, welchen Standards zukünftige Tests in Musiktheorie und Gehörbildung folgen sollten und inwieweit aus psychometrischer Sicht Verbesserungsbedarf an den aktuellen Tests besteht.

Der Berufsverband deutscher Psychologinnen und Psychologen hat die DIN-Norm 33430 zu diesen Fragen entwickelt (Hornke, 2005; Kersting, 2008), die einen Beitrag zur Umsetzung, Entwicklung und Verbesserung von Prozeduren zur Eignungsfeststellung leisten will. Eine musikpraktische und -theoretische Zulassungsprüfung ist demzufolge, obwohl sie der Berufsausbildung noch vorgeschaltet ist, im klassischen Sinne eine Eignungsbeurteilung. Sie testet einen Teilbereich der edukativen/biografischen, psychologischen und medizinischen Merkmale des Bewerbers (Hornke, 2005), wovon erstere an allen drei der in Abschnitt 1.1

analysierten Hochschulen getestet werden und die psychologischen Merkmale an der Hochschule für Künste in Bremen im Rahmen des Eignungsgesprächs.

Die Norm definiert Kriterien und –standards für Verfahrensweisen zur berufsbezogenen Eignungsbeurteilung, für die Auswahl und Zusammenstellung von Aufgaben (darunter die drei Hauptgütekriterien auf Objektivität, Reliabilität und Validität), sowie für die Durchführung und Auswertung. Die Entwickler eines solchen Tests benötigen auch gewisse Voraussetzungen um einen Test zu entwerfen, durchzuführen und auszuwerten, wobei neben dem Wissen um die zu prüfenden Inhalte auch Kenntnisse zu möglichen Testformen und statistischen Analyseinstrumenten, wie z. B. der Testtheorie, erforderlich sind. Die hier genannten Gütekriterien eines Testinstrumentes werden mit Ausnahme der Skalierbarkeit (nur in der DIN-Norm) von Bühner (2011) bestätigt. Kersting (2008) stellt im Anhang eine Checkliste mit den Inhalten der DIN-Norm vor, womit festgestellt werden kann, in welchen Bereichen der geprüfte Test von der Norm abweicht.

1.4 Zielsetzung der Studie

Im Folgenden wird eine statistische Analyse von Daten aus einer Zulassungsprüfung in Musiktheorie und Gehörbildung in drei Schritten durchgeführt, um auf dieser Grundlage eine der verwendeten Prüfungsaufgaben (Akkordbenennungsaufgabe) exemplarisch zu evaluieren.

Erstens soll untersucht werden, ob die untersuchte Prüfung der DIN-Norm entspricht und inwieweit die erzielten Ergebnisse der Teilnehmer mit theoretisch angenommenen Verteilungsfunktionen übereinstimmen.

Zweitens sollen die Lösungswahrscheinlichkeiten der Items einer Akkordbenennungsaufgabe mit den Mitteln der klassischen Testtheorie bestimmt werden, und es wird geprüft, ob diese mit der Normalverteilungsannahme übereinstimmen.

Drittens wird der Item-Fit anhand des Rasch-Modells (1PL-Modell) berechnet, um etwaige Items auszuschließen, die nicht modellkonform sind. In einem weiteren Schritt wird dann die Modellgüte des Rasch-Modells bestimmt und überprüft, ob die in der Zulassungsprüfung verwendeten Items eine sinnvolle Bandbreite von leichten und schwierigen Items darstellen, was wiederum eine Voraussetzung für eine differenzierte Leistungserfassung von Bewerbern ist.

2 Methode

2.1 Design

Datengrundlage für die testtheoretische Analyse waren im ersten Schritt die Ergebnisse einer schriftlichen Zulassungsprüfung im Fach Musiktheorie für einen lehramtsbezogenen Studiengang an einer deutschen Musikhochschule. Insgesamt umfasste die Stichprobe drei Jahrgänge (2009-2011). In dieser Zeit bestand dieser Teil der Zulassungsprüfung jeweils aus unterschiedlichen Aufgaben, obwohl der gesamte Prüfungsteil immer auf einer einheitlichen Skala von 0 bis 15 Punkten bewertet wurde.

Im zweiten Schritt wurde eine Akkordbenennungsaufgabe aus der Zulassungsprüfung lehramtsbezogener Studiengänge des Jahres 2011 ausgewertet. Die insgesamt 30 Items (zwei Klausurversionen mit je 15 Items, siehe Appendix A) bestanden aus je einem Akkord, der in einem offenen Antwortformat von den Prüflingen in einem von ihnen präferierten musiktheoretischen Benennungssystem bezeichnet werden sollte.

Sowohl die Punktzahl jedes Bewerbers der schriftlichen Musiktheorie-Teilprüfung als auch die Antwortmuster der Akkordbenennungsaufgabe wurden dem Archiv der Hochschule entnommen und anonymisiert, wobei das Präsidium der Hochschule die erforderliche datenrechtliche Genehmigung erteilte. Anschließend wurden die überprüften Benotungsverteilungen aller drei Jahrgänge sowie die Antwortmuster der Teilaufgabe des Jahres 2011 auf der Ebene der einzelnen Items ausgewertet. Die 30 Akkordbestimmungsaufgaben wurden von zwei Experten (Musikwissenschaftlern) als Musterlösung benannt und zu einer der Korrektur dienenden Musterlösung zusammengefügt. Die Antworten der Bewerber wurden hiermit verglichen. Der Fall von nicht einheitlichen Lösungsvorschlägen trat nur bei mehrdeutigen Akkorden auf (z. B. kleine Moll-Septakkorde, die auch als Dur Sixte ajoutée Akkorde interpretiert werden können), wobei nur solche Lösungen, die von beiden Experten anerkannt wurden, Aufnahme in die Bewertungsschablone fanden.

2.2 Stichprobe

Im Jahr 2009 haben $n_1 = 84$ Personen an der Zulassungsprüfung für den lehramtsbezogenen fächerübergreifenden Bachelor teilgenommen, 2010 waren es $n_2 = 86$ Personen und 2011 $n_3 = 132$ Personen. Anhand dieser Bewerberzahlen wurden die Verteilungen der Punkte berechnet. Von den $n_3 = 132$ Teilnehmern aus dem Jahr 2011 konnten zusätzlich die Ergebnisse einer Akkordbenennungs-Teilaufgabe von $n_{3a} = 124$ ausgewertet werden. Von diesen 124 Teilnehmern haben 64 (52%) die Prüfung bestanden. Von den zwei Versionen der Akkordbenennungsaufgabe mit je 15 Akkorden erhielten 73 Teilnehmer (59%) die Version 1 und 51 Teilnehmer (41%) die Version 2.

2.3 Aufgaben

Die untersuchte Klausur in Musiktheorie bestand im Jahr 2011 im ersten Teil aus einem Gehörbildungsteil mit einem melodisch-rhythmischen Diktat und einer Aufgabe zum Notieren vorgespielter Akkorde. Im zweiten Teil der Prüfung wurden anhand einer Akkordbenennungsaufgabe mit 15 Items und einer fünftaktigen Harmonisierung zu einem gegebenen Bassverlauf Kenntnisse in Musiktheorie getestet. Im dritten Teil wurden Musikstücke vorgespielt, zu denen in einem offenen Antwortformat Kontextwissen geprüft wurde (z. B. Taktart, Tempo, Gattung, Epoche). Die Bearbeitungsdauer für die gesamte Prüfung betrug eine Stunde.

Sowohl für die Höraufgaben als auch für die Harmonisierungsaufgabe ist eine testtheoretische Analyse praktisch unmöglich, da die Aufgabenstellung einer solchen Aufgabe nicht die Kompositionstechniken einer bestimmten Epoche verlangt und auch kein Konsens innerhalb der Musiktheorie über ein allgemein verbindliches Regelwerk für solche Aufgaben herrscht (de la Motte, 2004). Außerdem ist die Handhabung von Folgefehlern sowie der genaue Punkteerwerb oder -verlust nicht objektiv ermittelbar, da eine Vielzahl korrekter Lösungen möglich sind (siehe Weidner, 2010). Die objektiv richtigen Lösungen sind für eine Aufgabe, deren Aufgabenstamm „Harmonisieren Sie den vorgegebenen Bassverlauf mit einem sinnvollen Akkordsatz“ oder „Satzaufgabe“ (HMTMM, o. J.) lautet, im Nachhinein nicht mehr zu bestimmen, da diese zum Zeitpunkt der Prüfung weder definiert noch im Aufgabenstamm klar kommuniziert war.

Die Akkordbenennungsaufgabe ist für eine solche statistische Reanalyse hingegen gut geeignet, da eine oder mehrere richtige Lösungen existieren und anhand dieser der Lösungsvorschlag der Kandidaten als eindeutig richtig oder falsch klassifiziert werden kann. In dieser Benennungsaufgabe wurden die folgenden Akkorde getestet: Dur-Dreiklang, Moll-Dreiklang, verminderter Dreiklang, übermäßiger Dreiklang, kleiner und großer Dur-Septakkord, kleiner und großer Moll-Septakkord sowie halb- und vollverminderter Septakkord in verschiedenen Umkehrungen (siehe Appendix A).

Die Form der Beantwortung wurde den Teilnehmern freigestellt, sodass Bewerber mit den Studien-Schwerpunkten Klassik bzw. Jazz, Rock und Pop die jeweils ihnen vertrauten Beschreibungssysteme (Funktions- bzw. Stufentheorie) verwenden konnten. In der Nachkorrektur wurde dementsprechend auch das vom jeweiligen Bewerber verwendete System äquivalent zur „klassischen“ (funktionstheoretischen) Lösung angewendet.

3 Ergebnisse

Zunächst ist zu klären, in welchen Punkten die Zulassungsklausur der DIN-Norm nicht entspricht. Zu diesem Zweck bietet sich die Checkliste 1 „Anforderungen an Verfahrenshinweise, Normative Aussagen“ des DIN Screen (Kersting, 2008) an, die als sogenanntes „Negativ-Screening“ geeignet ist, um Tests auf Konformität mit der DIN-Norm zu prüfen.

Die erste Aussage muss bereits verneint werden: „Für jedes Verfahren liegen Verfahrenshinweise (Handhabungshinweise) vor“ (Kersting, 2008, S. 144). Hierunter fällt für einen Test wie diesen ein Manual, in dem u. a. die korrekten Lösungsantworten enthalten sein müssten. Mit der Nichterfüllung dieser Anforderung ist die gesamte Checkliste nicht erfüllt und somit die Zulassungsklausur nicht DIN-konform, da keine geregelten Hinweise zur Handhabung existieren.

Nichtsdestoweniger lassen sich Aussagen über die Fertigkeiten der Bewerber und die Eigenschaften der Aufgaben treffen, da ein Test wie diese Zulassungsklausur einerseits in Zukunft problemlos mit vorab festgelegten Verfahrenshinweisen durchgeführt werden kann und andererseits sämtliche Aufgaben von Experten korrigiert wurden, die teilweise die Klausur auch entworfen haben.

3.1 Theoretische Annahmen über die Klausur nach der klassischen Testtheorie

Für diese Verteilung wird eine große Anzahl der Werte beim gewählten Cutoff-Wert, also bei einem Mittelwert bzw. Median von etwa $M = 7$ erwartet, um die maximal mögliche Differenzierung der Klausurergebnisse gerade in dem Bereich, der über Prüfungserfolg oder -misserfolg entscheidet, zu erfassen. Wegen dieser Konstruktion und der Annahme, dass organisatorische und finanzielle Hürden (z. B. Anreise, Entrichten einer Verwaltungsgebühr) nur wenige schlecht vorbereitete Bewerber teilnehmen werden und dass extrem gute Leistungen in diesem schwierigen Test sehr unwahrscheinlich sind, wird eine Normalverteilung angenommen.

In der nun folgenden Auswertung nach der klassischen Testtheorie bedienen wir uns einer reinen Messfehlertheorie, die den sogenannten „wahren“ Testwert einer Person aus den gemessenen Werten berechnet und somit eine Kontrolle der unsystematischen Schwankungen ermöglicht (siehe Bühner, 2011).

3.2 Verteilungen der Punkte in den Klausuren der Jahre 2009 bis 2011

Die aggregierte Verteilung der Gesamtpunktzahlen aus den Jahrgängen 2009 bis 2011 erfüllt nicht die Annahme der Normalverteilung nach dem Shapiro-Wilk Test ($W = 0,94$; $n = 302$; $p < 0,001$) (Appendix B, Abb. 1). Der Mittelwert von $M_{agg} = 5,06$ ist deutlich niedriger als der erwartete von $M = 7$ (Wilcoxon-Vorzeichen-Rangtest, $V = 9513,5$; $p < 0,001$), der $Median_{agg} = 4$ liegt weit links von der Skalenmitte.

In der schriftlichen Zulassungsprüfung Musiktheorie des Jahres 2009 (Appendix B, Abb. 2) ist der Mittelwert $M_{2009} = 4,73$, der $Median_{2009} = 3,0$ und die Verteilung ist bimodal mit Modi bei 2,0 und 10,5 Punkten (nach Shapiro-Wilk: $W = 0,88$; $n = 84$; $p < 0,001$). Im Jahr 2010 (Appendix B, Abb. 3) liegt der Mittelwert bei $M_{2010} = 4,36$, der Median wieder bei $Median_{2010} = 3,0$ und auch diese Verteilung ist bimodal (Modi bei 1,0 und 11,0 Punkten) (nach

Shapiro-Wilk: $W = 0,89$; $n = 86$; $p < 0,001$). Für das Jahr 2011 (Appendix B, Abb. 4) liegen der Mittelwert von $M_{2011} = 5,73$ und der Median von $Median_{2011} = 5$ deutlich näher beieinander, allerdings ist auch diese Punkteverteilung nicht normalverteilt (nach Shapiro-Wilk: $W = 0,96$, $n = 132$, $p = 0,002$). Wie in den Jahren 2009 und 2010 ist auch die Punkteverteilung der Prüfung des Jahres 2011 rechtsschief. In Tab. 3 sind diese Ergebnisse dargestellt, wobei sowohl parametrische als auch nicht-parametrische Kennwerte angegeben sind, die gerade in ihren Unterschieden noch einmal deutlich machen, dass die Annahme der Normalverteilung verletzt ist. Hervorzuheben ist auch der Interquartilsabstand, dass sich also 50% der Bewerber in einem eher schmalen Punktebereich (5 oder 6 Punkte) befinden und somit eine eher schmale Streuung vorliegt.

Tab. 3: Überblick über die statistischen Kennwerte der Verteilungen der Klausurergebnisse von 2009 bis 2011

	2009	2010	2011	Gesamt (2009-2011)
n	84	86	132	302
Mittelwert	4,73	4,36	5,73	5,06
Standardabweichung	3,89	3,59	3,27	3,58
Median	3	3	5	4
Interquartilsabstand (QA = $Q_{75} - Q_{25}$)	6	5	5	6
Spannweite	13	12	14	14
Schiefte	0,734	0,648	0,482	0,547
Kurtosis	-0,765	-0,817	-0,296	-0,681

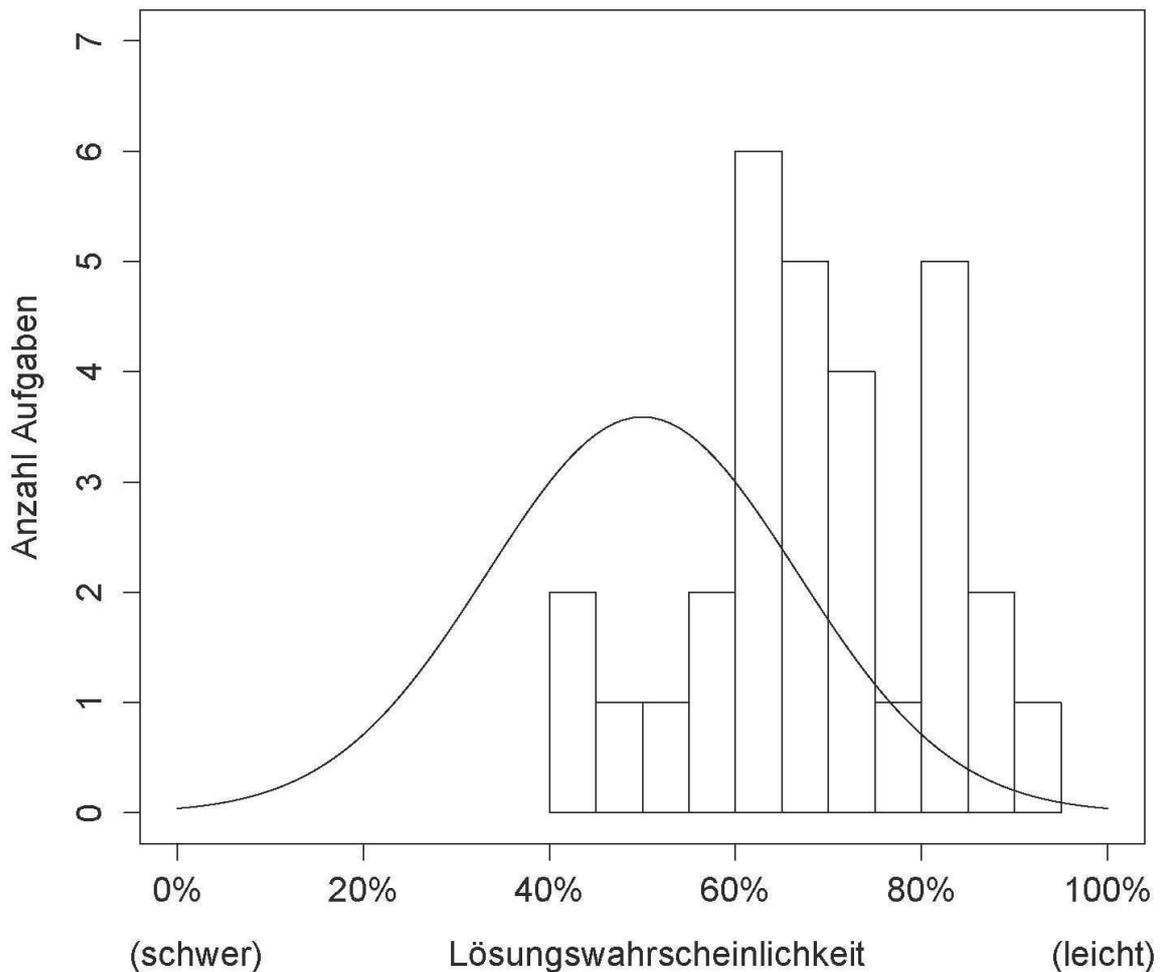
Die je nach Jahr verschiedene Variabilität und Schiefe der Ergebnisverteilungen könnte sowohl auf unterschiedlich schwierige Aufgaben als auch auf unterschiedlich leistungsfähige Bewerber hinweisen. Mit der klassischen Testtheorie ist diese doppelte Erklärung der Ergebnisunterschiede nicht auflösbar. Aus diesem Grund sollen im letzten Analyseschritt die Prüfungsergebnisse als Zusammenspiel von Bewerberleistung und Aufgabenschwierigkeit mit Hilfe der probabilistischen Testtheorie transparent gemacht werden.

3.3 Lösungswahrscheinlichkeiten der Akkordbenennungsaufgabe und der einzelnen Items (2011)

In diesem nächsten Schritt werden wir eine einzelne Aufgabe zur Akkordbenennung untersuchen, weil Unregelmäßigkeiten für die gesamte Prüfung festgestellt wurden. Da kein signifikanter Gruppenunterschied der Lösungswahrscheinlichkeiten zwischen den beiden Klausurversionen vorlag (Wilcoxon-Mann-Whitney-Test, $W = 67$; $p = 0,06$), wurden die Daten der beiden Aufgabenversionen zusammengefasst.

Gemessen an der durchschnittlich erreichten Gesamtpunktzahl der Zulassungsprüfung (5,7 Punkte) ist die Akkordbenennungsaufgabe eine eher leichte Aufgabengruppe innerhalb der Klausur (mittlere Lösungswahrscheinlichkeit: $P = 0,69$), wobei die leichteste Aufgabe von 94% und die schwerste noch von 41% aller Bewerber gelöst wurden. Die Verteilung der Lösungswahrscheinlichkeiten der Aufgaben (s. Abb. 1 und Appendix A) ist normalverteilt (nach Shapiro-Wilk: $W = 0,97$; $n = 30$; $p = 0,59$), jedoch stark linksschief im Vergleich zu einer Normalverteilung mit Mittelwert von $P = 0,5$ ($v = -0,30$; $t(29) = 7,65$; $p < 0,001$; $d = 1,40$; $CI\ 95\% = [0,91;1,88]$).

Abb. 1: Verteilung der Lösungswahrscheinlichkeiten für die 30 Akkordbestimmungsaufgaben aus der Musiktheorieklausur des Jahres 2011.



3.4 Probabilistische Testtheorie und das Rasch-Modell

Innerhalb der klassischen Testtheorie bleiben Fragen nach der Validität der Items und der gesamten Aufgabe unbeantwortet. Außerdem berücksichtigt die klassische Theorie nur, dass zwei Bewerber eine gleiche Anzahl von Aufgaben richtig beantwortet haben, jedoch nicht, ob die gelösten Aufgaben leicht oder schwer waren, d. h. also, ob also die *Aufgaben* für gute Bewerber zu schwer, oder ob die *Bewerber* trotz durchaus lösbarer Aufgaben zu schlecht waren.

Eine Lösung für diese Anforderungen bietet die folgende Analyse: Mit der probabilistischen Testtheorie (engl. *Item response theory*) kann die Schwierigkeit jedes messbaren Items innerhalb eines Tests und der latenten Variable, wie hier der Personenfähigkeit⁶, bestimmt

⁶ Im methodischen Teil wird – im Gegensatz zur bisherigen Terminologie – vorwiegend von „Personenfähigkeiten“ gesprochen, da dies in der Testtheorie so üblich ist. Es sind jedoch weiterhin bereits erworbene und messbare Fertigkeiten in Musiktheorie bzw. Gehörbildung gemeint.

werden. Um diese beiden Informationen zu messen, wird angenommen, dass die Items statistisch unabhängig voneinander sind und ihr Zusammenhang nur von der latenten Variablen abhängt. Der hier angenommene Spezialfall der Testmodelle, das Rasch-Modell, basiert auf der Annahme, dass die gesamte Information über die Personenvariable im beobachteten Muster der Itemantworten liegt. Mithilfe der Maximum-Likelihood-Methode wird die Wahrscheinlichkeit dafür berechnet, dass die beobachteten Daten unter der Annahme gemessen worden sind, dass das angenommene Modell gilt. Die Likelihood-Methode ist also ein Maß zur Einschätzung der Güte oder Mutmaßlichkeit des Modells. Im Zuge dessen werden auch die Parameter für die Itemschwierigkeit und die Personenfähigkeit geschätzt.

Im Rasch-Modell definieren die beiden Parameter Itemschwierigkeit (σ) und Personenfähigkeit (θ) die Lösungswahrscheinlichkeit des Items, wobei die Lösungswahrscheinlichkeit für ein Item bei $P = 0,5$ liegt, wenn Personenfähigkeit und Itemschwierigkeit gleich groß sind (siehe Formel 1, nach Rost, 2004).

$$P(X_{is} = 1 | \theta_s, \sigma_i) = \frac{e^{(\theta_s - \sigma_i)}}{1 + e^{(\theta_s - \sigma_i)}}$$

(Formel 1)

3.5 Analyse der Testdaten mittels des Rasch-Modells (1PL)

Zunächst soll die Itemschwierigkeit σ kontrolliert werden, die mit der Personenfähigkeit θ die Berechnungsgrundlage des Rasch-Modells liefert (die Berechnung erfolgte mit dem Paket eRm in R; siehe Mair, Hatzinger & Maier, 2011). In einem Kompetenztest, der als Grundmodell für die Zulassungsprüfung angenommen wird, sollte idealerweise das Intervall $\sigma = [-3; 3]$ der Itemschwierigkeit durch die berechneten Logit-Kennwerte⁷ annähernd ausgeschöpft werden (Bond & Fox, 2007). In den Daten der Zulassungsprüfung liegen die Schwierigkeiten für die Aufgabenversion 1 jedoch bei $\sigma_1 = [-1,2; 1,3]$ (siehe Abb. 2 und Appendix C1) und in der Version 2 in einem etwas weiteren Bereich von $\sigma_2 = [-2,2; 2,0]$ (siehe Abb. 3 und Appendix C2), wobei die beiden Items, die das Intervall eingrenzen, nur von je einer einzigen Aufgabe vertreten werden. In beiden Versionen decken die Aufgaben also nur einen äußerst begrenzten Bereich von Itemschwierigkeiten ab, den wir aufgrund der klassischen Analyse dem eher leichten Spektrum zuordnen können. Letzteres trifft insbesondere auf die Aufgabenversion 2 zu, welche die vier Aufgaben mit den höchsten Lösungswahrscheinlichkeiten $P = \{94\%, 86\%, 86\%, 84\%\}$ enthält (siehe Appendix A). Diese Aufgabe,

⁷ Der Logit-Wert einer Wahrscheinlichkeit entspricht dem natürlichen Logarithmus des Bruches der Lösungswahrscheinlichkeit geteilt durch die entsprechende Gegenwahrscheinlichkeit, z. B. $\ln(0,8/0,2) = 1,39$ für eine Lösungswahrscheinlichkeit von 80%. Diese Logit-Werte können grundsätzlich zwischen $-\infty$ und $+\infty$ liegen, befinden sich jedoch in der Regel zwischen $[-3;3]$.

die als einzige in ihrer Form statistisch auswertbar ist, erweist sich also als ungeeignet, um das vorhandene Fertigkeitsspektrum der Bewerber umfassend zu prüfen.

Im oberen Bereich der Abbildungen 2 und 3 befindet sich die jeweilige Verteilung der Personenfähigkeiten der Bewerber, ein weiterer Hinweis darauf, dass die Fähigkeitsverteilung für diese Aufgabe linksschief ist, die Aufgaben also eher leicht sind.

Abb. 2: Person-Item Map für die Aufgaben der Klausurversion 1 als Darstellung der Itemschwierigkeit (Items 1-15 auf der Ordinate eingezeichnet) und der Personenfähigkeit nach der Item-Response-Theorie.

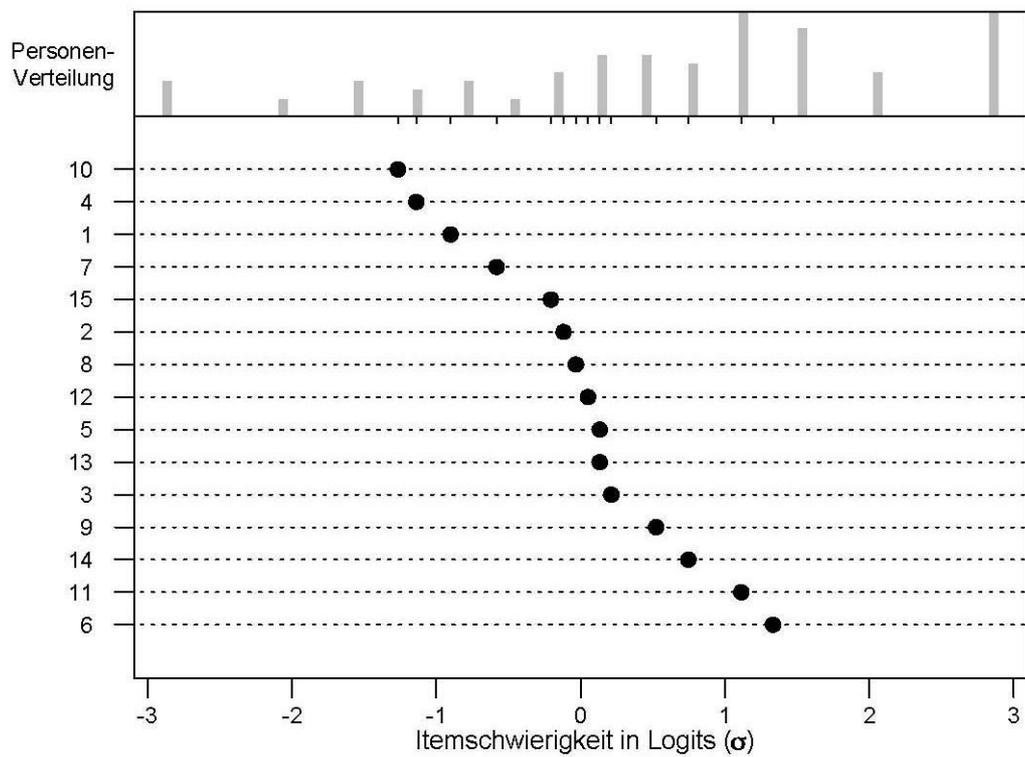
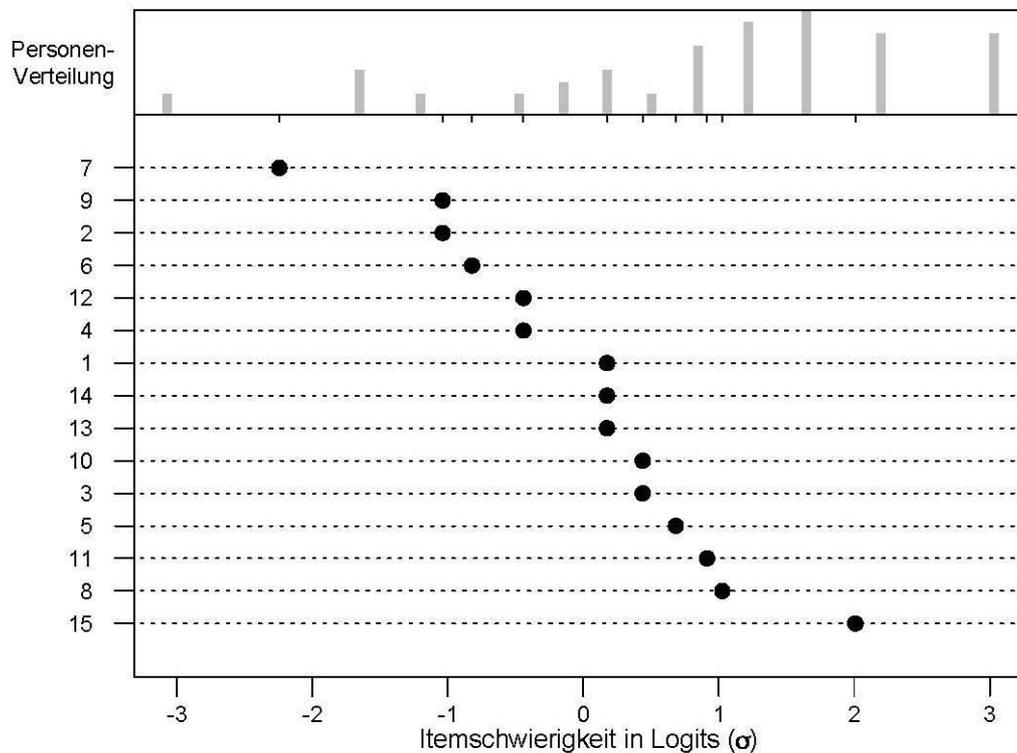


Abb. 3: Person-Item Map für die Aufgaben der Klausurversion 2 als Darstellung der Itemschwierigkeit (Items 1-15 auf der Ordinate eingezeichnet) und der Personenfähigkeit nach der Item-Response-Theorie.



3.5.1 Modellgeltungstests zur Überprüfung des Rasch-Modells (1PL)

Zur Überprüfung der Modellgüte wurde der Likelihood-Quotienten-Test (LR-Test) nach Andersen (1973) durchgeführt. Dieser Test beruht darauf, dass das Rasch-Modell gilt, wenn sich die Aufgaben-Parameter in verschiedenen Subgruppen nicht voneinander unterscheiden. Hier wurde ein randomisiertes Splitkriterium angenommen, da in den Subgruppen für gewisse Items nur richtige oder falsche Antwortlösungen vorlagen. Um zufällig signifikante oder nicht-signifikante Ergebnisse möglichst auszuschließen, wurde der Andersen-Test mit 200 randomisierten Splitkriterien-Vektoren berechnet. Maße der zentralen Tendenz konnten generell bei Aufgaben mit hohen Lösungswahrscheinlichkeiten nicht als gruppenbildende Kriterien angewendet werden können, da die Antwortmuster dieser Aufgaben nicht in verschiedene Personen-Gruppen unterteilt werden konnten; in diesen Fällen sind randomisierte Splitkriterien eine Möglichkeit, die folgenden Tests durchführen.

Bei der Modellberechnung für Klausurversion 1 waren die Subgruppen in 19 von 200 Modellen signifikant unterschiedlich ($p < 0,05$), was unter der Annahme einer Binomialverteilung nicht mehr innerhalb einer Standardabweichung liegt (Erwartungswert $E = 10$; $s = \text{Sqrt}[n \cdot p \cdot (1 - p)] = 7,1$; siehe Tab. 4), ein Rasch-Modell-Fit wurde somit vorerst abgelehnt.

Tab. 4: Der Andersen-LR -Test für Klausurversion 1

Klausurversion 1	LR-Wert	p	Log-likelihood	Log-likelihood
			“Splitkrit. = 0”	“Splitkrit. = 1”
Mittelwert	14,9	0,46	-199,1	-189,0
Minimum	3,78	0,01	-261,5	-253,1
Maximum	37,8	0,99	-127,2	-126,1

In der Klausurversion 2 wichen in neun von 200 Berechnungen die Subgruppen signifikant voneinander ab (Tab. 5), was noch unter dem Erwartungswert von $E = 10$ liegt, also kann das Rasch-Modell hier angenommen werden.

Tab. 5: Der Andersen-LR -Test für Klausurversion 2

Klausurversion 2	LR-Wert	p	Log-likelihood	Log-likelihood
			“Splitkrit. = 0”	“Splitkrit. = 1”
Mittelwert	14,2	0,48	-109,8	-111,9
Minimum	2,9	0,01	-167,7	-164,9
Maximum	30,3	0,99	-57,2	-51,0

3.5.2 Modellanalyse auf der Itemebene

Für die folgende Modellüberprüfung müssen die Items aus beiden Aufgabenversionen wieder getrennt werden. Hierzu wurde zunächst der Q-Index herangezogen, der ein Maß für den Item-Fit ist. Im Gegensatz zum z. B. MNSQ-Index schätzt der Q-Index die konditionale Likelihood ohne Informationen über Personenparameter, sodass der Items-Fit ausschließlich aufgrund der Itemeigenschaften bestimmt wird (Rost & von Davier, 1994). In keiner der beiden Aufgabenversionen musste ein Item aufgrund eines zu hohen oder eines zu niedrigen Q-Indexes ausgeschlossen werden, denn jeder Item-Fit lag im Bereich von $-1,96 < z(Q\text{-Index}) < 1,96$ (die Berechnung erfolgte mit der Software Winmira; s. Davier, 2001). Aufgrund des t-standardisierten Outfits und Infit Mean Squares wurden die Items 1 und 3 aus Klausurversion 1 und das Item 1 aus Klausurversion 2 herausgenommen (siehe Appendix C1 und C2).

Da auch der Wald-Test auf dem Vergleich von Subgruppen beruht, wurden hier 1500 Berechnungen mit randomisierten Splitkriterien durchgeführt. Ab 94 signifikanten Abweichungen ($E + 1 \cdot s \geq 94$) wurde ein Item überzufällig häufig als modellunpassend befunden ($E = 75; s = \text{Sqrt}[n \cdot p \cdot (1 - p)] = 19,3$). Dies wurde angenommen, da den Splitkriterien keine aussagekräftigen Parameter (wie z. B. Geschlecht) zugrunde liegen und die möglicherweise systematischen Unterschiede in den Subgruppen nicht in jeder Kombination von Splitkriterien auftreten.

Folglich wurden in Klausurversion 1 insgesamt drei Items (Akkorde 1, 3 und 9) und in Klausurversion 2 insgesamt zwei Items (Akkorde 1 und 5) für die Modellanalyse eliminiert (siehe Appendix A und C1 und C2, ausgeschlossene Items sind kursiv gedruckt).

Eine Wiederholung des Andersen-Tests mit den verringerten Itemanzahlen führte nun zu 12 (Klausurversion 1) bzw. 13 (Klausurversion 2) signifikanten Abweichungen (Tab. 6 & 7), wobei der maximale Erwartungswert von $E = 17$ signifikanten Ergebnissen (siehe 3.5.1.) nicht überschritten wird und wir das Rasch-Modell für die reduzierten Modelle weiterhin annehmen können.

Tab. 6: Der Andersen-LR-Test für Klausurversion 1 (12 Items)

Klausurversion 1	LR-Wert	p	Log-likelihood	Log-likelihood
			“Splitkrit. = 0”	“Splitkrit. = 1”
Mittelwert	11,1	0,50	-137,5	-138,2
Minimum	2,12	0,00	-192,6	-187,0
Maximum	28,8	1,00	-87,4	-81,9

Tab. 7: Der Andersen-LR-Test für Klausurversion 2 (13 Items)

Klausurversion 2	LR-Wert	p	Log-likelihood	Log-likelihood
			“Splitkrit. = 0”	“Splitkrit. = 1”
Mittelwert	13,1	0,44	-84,5	-84,7
Minimum	2,80	0,00	-123,9	-123,7
Maximum	34,9	1,00	-42,8	-47,6

Nach der Eliminierung der nicht konformen Items wird das Rasch-Modell für beide Klausurversionen angenommen.

Zum Vergleich der ursprünglichen mit den reduzierten Modellen werden die konditionalen Loglikelihood-Werte der beiden ursprünglichen Modelle mit den jeweils um zwei bzw. drei Items reduzierten Modellen verglichen. Durch die Item-Eliminierung und die Rasch-Modell-konforme Anpassung beider Modelle verbessert sich die Likelihood in beiden Versionen der Klausur. In Klausurversion 1 mit allen 15 Testitems beträgt die Loglikelihood $L_{1alle} = -395,6$; werden die Items 1, 3 und 9 eliminiert, beträgt die Loglikelihood $L_{1elim} = -295,0$. In Klausurversion 2 mit allen 15 Testitems beträgt die Likelihood $L_{2alle} = -228,8$; werden die Items 1 und 5 eliminiert, beträgt die Likelihood $L_{2elim} = -175,7$. Die Steigerung der konditionalen Loglikelihood-Werte innerhalb der jeweiligen Klausurversionen ist ein Indikator für eine deskriptive Verbesserung beider nun Rasch-konformen Modelle.

4 Diskussion

Ziel dieser Studie war es, eine musiktheoretische Zulassungsklausur zu einem musikpädagogischen Studiengang mit psychometrischen und testtheoretischen Voraussetzungen zu vergleichen. Des Weiteren sollte mittels der klassischen Testtheorie eine exemplarische Klausuraufgabe (Akkordbenennungsaufgabe) auf Normalverteilung überprüft werden. Zuletzt wurde diese Aufgabe unter Verwendung der probabilistischen Testtheorie auf Item-Fit und Modellkonformität überprüft. Es sollte ermittelt werden, ob die verwendeten Items die Fertigkeit der Akkordbenennung als Teil der Musiktheorie-Kompetenz umfassend darstellen können. Vor dem Hintergrund unserer Ergebnisse bleibt für die musiktheoretische Zulassungsprüfung zu prüfen, ob nicht die ermittelte Ergebnisverteilung Anlass für eine testtheoretisch gestützte Neukonzeption unter Einbeziehung eines Kompetenzmodells für Musiktheorie und Gehörbildung sein sollte. Diese Neukonzeption sollte die bereits definierten Standards für allgemeine Eignungsbeurteilungen einbeziehen, die in die DIN-Norm 33430 Eingang gefunden haben, da diese Standards bisher nicht eingehalten werden (Hornke, 2005).

Ein Hinweis auf die Notwendigkeit einer solchen konzeptionellen Verbesserung ist bereits die Punkteverteilung der Klausuren der Jahre 2009 bis 2011. Die durchschnittlich erzielte Punktzahl der letzten drei Jahrgänge lag immer deutlich unter dem vorher festgelegten Cutoff-Wert von 7 Punkten. Die daraus resultierende negative Differenz zwischen dem tatsächlichen und dem erwarteten Mittelwert wirft die Frage auf, ob die Aufgaben der Prüfung zu schwer sind, die Teilnehmer zu schlecht vorbereitet oder ob es möglicherweise sogar einen ganz anderen Grund dafür gibt, dass die Notenverteilung so rechtsschief ausfällt. Möglich ist jedoch auch, dass gerade eine schiefe oder bimodale Verteilung erzielt werden soll, dass also genau eine solche klare Trennung der "besseren" und "schlechteren" Bewerber erfolgen soll, was durch eine bimodale Punkteverteilung dargestellt werden kann. Solch eine durchaus diskutabile Konzeption der Zulassungsprüfung müsste allerdings theoretisch gestützt, im Vorhinein definiert werden und vor allem in mehreren Jahrgängen konsistent auftreten, was in unseren drei Stichproben jedoch nicht der Fall ist.

Dementsprechend nehmen wir weiterhin an, dass ein musiktheoretischer Fertigkeitstest im Allgemeinen zu normalverteilten Ergebnissen führen soll. Er soll eine kleine Spitzengruppe identifizieren. Die Anforderungen dieses Tests sollen außerdem transparent kommuniziert werden, sodass im Idealfall nur die Bewerber an der Zulassungsprüfung teilnehmen, die Studierfähigkeit vorweisen und die Klausur (wahrscheinlich) bestehen können. Die meisten Bewerber werden gut vorbereitet sein, da sie wissen, welche Fertigkeiten gefordert sind, sie werden bürokratische Hürden genommen haben und somit eine mittlere Note erzielen. Diese Annahmen führen idealerweise zu insgesamt normalverteilten Klausurergebnissen. Aufgrund dieser theoretischen Überlegungen zu einer erwünschten Normalverteilung der Klausurergebnisse gehen wir davon aus, dass den Klausuren der Jahre 2009 und 2010 vor allem mittelschwere Items fehlen und die Testergebnisse deshalb eine Lücke im Skalenbereich von 7 Punkten aufweisen. Im Jahr 2011 wird dies durch die überdurchschnittlich leichte Aufgabe zur Akkordbenennung ausgeglichen. Ohne eine genauere Steuerung mittels besserer Kenntnis der Itemqualität und einer für die Aufnahmeprüfungskandidaten transparenten Diskussion der Ziele der musiktheoretischen Zulassungsprüfung kann die Frage nach einer zielgerichteten Konzeption der Leistungsmessung nicht beantwortet werden.

Einige der in der musiktheoretischen Zulassungsprüfung von 2011 verwendeten und hier nicht weiter analysierbaren Klausuraufgaben versuchen durchaus anspruchsvolle Fertigkeiten zu testen. Durch das jeweilige Aufgabendesign können jedoch in diesen Fällen keine Aussagen über die Validität der Items getroffen werden. Es bleibt letztlich unklar, was die Items eigentlich messen, nicht zuletzt durch uneindeutige Aufgabenstellungen und generell fehlende Objektivität. Dies stellt ein zusätzliches Erschwernis für die Bewerber dar, da selbst bei einer Vorbereitung mit der hochschuleigenen Musterklausur nicht nachvollziehbar ist, welche Aufgabenlösung eine hohe Punktzahl erzielen würde⁸. Für andere Anforderungsbereiche, die in messtheoretische Items ähnlich der vorliegenden Akkordbenennungsaufgabe unterteilt werden können, verläuft die Zusammenstellung unkontrolliert: In der hier analysierten Akkordbenennungsaufgabe wird das Spektrum von leichten, mittleren und schweren Aufgaben nicht zweckmäßig ausgeschöpft; die Verteilung ihrer Lösungswahrscheinlichkeiten ($n = 30$) ist zwar noch normalverteilt, aber stark linksschief. Das bedeutet, dass sehr viele leichte und keine schwierige Items vorliegen. Demgegenüber sollten insbesondere schwierige Aufgaben innerhalb der Akkordbenennungsaufgabe von großem Interesse sein, da sie die Leistungsfähigkeit der Spitzengruppe und deren erworbene Fertigkeiten differenzierter erfassen.

⁸ In Großbritannien hat das Associated Board of the Royal Schools of Music umfangreiches und aktuelles Lehr- und Testmaterial für acht Klassenstufen in Musiktheorie und Gehörbildung entwickelt, die objektiv nach vorher festgesteckten Kriterien bewertet werden. Diese Testhefte sind ein Beispiel dafür, dass anspruchsvolle Aufgaben in diesen Bereichen durchaus objektiv gestellt und bewertet werden können. Ein Überblick über die Inhalte gibt folgende Webseite: <http://www.abrsm.org/en/exams/gradedMusicExams/latestSyllabuses.html> [12.08.12]

Für beide Klausurversionen der Akkordbenennungsaufgabe konnte das Rasch-Modell nach dem Eliminieren unpassender Items bestätigt werden, und zwar mittels des Likelihood-Quotienten-Tests (Andersen, 1973) und des Wald-Tests (Strobl, 2010). Die Itemanalyse beider Klausurversionen erbrachte für insgesamt fünf der 30 Items (siehe Appendix C1 und C2), dass Aufgabenparameter nicht mit den Parametern des Tests übereinstimmen. Wurden diese Items aus den Aufgaben entfernt, verbesserte sich der Modell-Fit in beiden Fällen deutlich. Grundsätzlich ist eine solche Aufgabe also geeignet, in einen umfassenderen Kompetenztest einbezogen zu werden, da sie eine eindimensionale Fertigkeit testet. Es ist nun zu prüfen, ob diese Fertigkeit zur Akkordbenennung in eine weiter gefasste musiktheoretische Kompetenzprüfung einbezogen werden kann.

Der mangelnde Item-Fit der ausgeschlossenen Items hat auch gezeigt, dass diese Items noch einmal unter Einbeziehung der zu erfassenden Gruppenmerkmale (z. B. Hauptinstrument oder stilistischer Schwerpunkt wie Klassik, Jazz oder Pop) getestet werden müssten, um die Einflüsse auf die Items, die im Wald-Test bedeutsam wurden, besser zu identifizieren. Diese Fragestellung lässt sich in einem ersten Schritt testen, indem die genannten Informationen in die Analyse von Musiktheorietests eingehen, also Instrumentengruppen zusammengefasst werden (z. B. Melodie- und Harmonieinstrumente, sowie transponierende Instrumente) und die genannten stilistischen Schwerpunkte erhoben werden. Die noch zu generierenden Hypothesen könnten anhand von Differential item functioning überprüft werden. Mit diesem Verfahren ließe sich klären, ob beispielsweise Pianisten tatsächlich bessere Leistungen im Hören von Harmoniefolgen erbringen⁹.

Eine solche Studie ist von besonderem Interesse, da Parameter wie das Hauptinstrument der Bewerber oder ihr stilistischer Schwerpunkt einen möglichen Einfluss auf das Ergebnis der (Musiktheorie-)Zulassungsprüfung haben können. Darauf deuten auch Untersuchungen zur Notenverteilung und prognostischen Validität der Zulassungsprüfung hin (Lehmann & Hornberger, 2010): Es hat sich gezeigt, dass Pianisten an der untersuchten Musikhochschule grundsätzlich bessere Note in Musiktheorie und Gehörbildung erreichen als andere Instrumentengruppen und Sänger. Ein weiteres Ergebnis waren große Unterschiede in der prognostischen Validität in den verschiedenen Instrumentengruppen beim Vergleich von Eignungs- und Diplomnote, wobei der Zusammenhang zwischen Eignungs- und Abschlussnote für Pianisten ($r = 0,59$) sehr viel besser ist als für Streicher ($r = 0,07$). Derartige Einflüsse des Hauptinstrumentes müssten mit den Daten anderer Hochschulen repliziert werden, um diese Effekte genauer zu spezifizieren.

⁹ Für unseren Datensatz war ein solches Vorgehen nicht möglich, da aus Gründen des Datenschutzes nur das Klausurblatt mit der Akkordbenennungsaufgabe verfügbar war.

5 Ausblick

In einem nächsten Schritt sollen Instrumente geschaffen werden, die diesen Status Quo an deutschen Musikhochschulen verbessern können. Um die aktuelle Situation für Lernende und Lehrende gleichermaßen zu verbessern, sollen ein Kompetenzmodell für Musiktheorie und Gehörbildung sowie dazugehörige Aufgaben entwickelt werden, die an einer für die Zulassungsprüfung repräsentativen Stichprobe empirisch zu überprüfen sind.

Zu diesem Zweck soll eine fachspezifische, strukturierte Repräsentation entwickelt werden, wie sie durch ein Kompetenzniveaumodell gegeben ist. Dieses Modell stellt in Abhängigkeit von bestimmten Zeitpunkten der Lernentwicklung Erwartungen an die Schüler dar und liefert „wissenschaftlich begründete Vorstellungen darüber, welche Abstufungen eine Kompetenz annehmen kann“ (Klieme et al., 2003). Insbesondere die verwendeten Aufgaben legen die Bedeutung der ausformulierten Kompetenzen fest, was auch den eingezeichneten Niveaugrenzen, die nur eine willkürliche Einteilung einer kontinuierlichen Variablen darstellen, entgegenwirkt (Hartig & Klieme, 2006).

Das Projekt zur genaueren Ausarbeitung des Kompetenzmodells und der Aufgabengenerierung erfolgt in Kooperation mit der Studienvorbereitenden Ausbildung (SVA) an niedersächsischen Musikschulen. Dabei werden der Schwierigkeitsgrad und die Gültigkeit von zu entwickelnden Aufgaben getestet und zu einem Itempool zusammengestellt, wobei zu berücksichtigen ist, dass die Aufgaben einerseits musiktheoretisch gültig als auch dem späteren Anforderungsbedarf eines Berufsmusikers entsprechen sollten (Rolle, 2008). Damit ist gemeint, dass die Stilkopie eines dodekaphonischen Liedes zwar eine objektiv bewertbare und musiktheoretisch valide Aufgabe ist, aber vermutlich nur wenige Schulmusiker diese Fertigkeit in ihrem späteren Berufsleben benötigen werden.

Zukünftige Testkonzeptionen könnten auf diesen validierten Items mit bekanntem Schwierigkeitsgrad aufbauen und somit auch die Personenfähigkeiten in konsekutiven Jahrgängen messen und vergleichen. Der Itempool könnte weiterhin eine theoretisch begründete Aufgabenerstellung ermöglichen, sodass Items bewusst eingesetzt und gesteuert werden können. Dadurch kann in Zukunft, z. B. mit adaptivem Testen, vermieden werden, dass Studierende Aufgaben lösen sollen, ohne die notwendigen Fertigkeiten zu besitzen (Karpinski, 2000; Estrada Rodriguez, 2008). Bei der Akkordbenennungsaufgabe ist offenkundig, dass insbesondere Aufgaben mit schwierigen Akkorden erstellt werden müssten. Die Regulierung der Itemschwierigkeit wäre bei der Akkordbenennungsaufgabe beispielsweise durch die Anzahl der Noten und Vorzeichen sowie durch die Verwendung von ungewohnten Schlüsseln und transponierenden Instrumenten als schwierigkeitsbestimmende Merkmale möglich. Diese Heuristik bedarf jedoch noch einer ausstehenden Prüfung.

Mit dem Ziel ein Kompetenzmodell und Aufgaben zu erfinden, zu testen und anschließend weiterzuentwickeln, wurde die Arbeitsgruppe „Kompetenzen und Aufgabenentwicklung für die Zulassungs-Teilprüfung Musiktheorie“ gegründet, die die Expertise von Fachleuten aus Musiktheorie und Testtheorie vereint und dadurch die formulierten Ziele langfristig verfolgen wird. Ein solches Kompetenzmodell kann über die gymnasialen Schulmusik-

studiengänge hinaus Bedeutung haben und in der gesamten Instrumentalpädagogik Anwendung finden. Das Modell wird zum ersten Mal die Integration von Musiktheorie und Gehörbildung in die instrumentalpädagogische Ausbildung auf der Grundlage empirischer Daten ermöglichen. Mit dieser generalisierbaren und zielführenden Facette der Instrumentalpädagogik kann dem Instrumentalschüler klar vermittelt werden, mit welchen Fertigkeiten sein Verständnis von Musik auf welche Art entwickelt und bereichert werden kann. Lehrenden steht ein Instrument in Form des Itempools zur Verfügung, mit dem klar definierte Kompetenzen getestet werden können und möglicherweise eine hochschulübergreifend vereinheitlichte Gehörbildungs-Zulassungsprüfung entwickelt werden kann, die an jeden Bewerber, unabhängig von seiner zuvor erlernten Gehörbildungsmethode, die gleichen Anforderungen stellt und im Vorhinein definiert, welche das sein werden (Estrada Rodriguez, 2008).

Sowohl den Vermittlern von Musiktheorie als auch den Aufnahmeprüfungskandidaten wäre somit eine bisher fehlende Orientierung geboten, die musiktheoretisch, testtheoretisch, pädagogisch und psychologisch bestätigt wurde. Diese transparente, objektive und valide Zielvorgabe kann schließlich das lebenslange wissende und emotional eingebundene Musizieren junger Menschen begünstigen und ihnen in ihrer Rolle als Musikvermittler behilflich sein.

Literaturverzeichnis

- Andersen, E. (1973). A goodness of fit for the Rasch model. *Psychometrika*, 38, 123-140.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, London: Lawrence Erlbaum Associates.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Davier von, M. (2001). *WINMIRA 2001 (Computersoftware)*. Groningen: Science Plus Group.
- De la Motte, D. (2004). *Harmonielehre*. Kassel u. a.: Bärenreiter
- De la Motte-Haber, H. (2005). Vorwort. In H. de la Motte-Haber & O. Schwab-Felisch, *Musiktheorie. Handbuch der Systematischen Musikwissenschaft, Band 2*. Laaber: Laaber-Verlag, S. 9-10.
- Estrada Rodriguez, L. A. (2008). *Didaktik und Curriculumentwicklung in der Gehörbildung. Eine vergleichende Untersuchung an deutschen Lehrbüchern zur Gehörbildung aus der Zeit 1889 bis 1985*. Hannover: Ifmpf.
- Hartig, J., & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hg.), *Leistung und Leistungsdiagnostik*. Heidelberg: Springer, S. 128-143.
- Hempel, C. (2008). *Neue Allgemeine Musiklehre*. Mainz: Schott.
- Hochschule für Künste Bremen (2007). *Aufnahmeprüfung Gehörbildung + Theorie*. Verfügbar unter: <http://www.musik.uni->

bre-
men.de/fileadmin/mediapool/musik/dateien/Aufnahmepruefung/HFK2007_MUSTERAU
FGABEN_MUSIK.pdf [12.09.12]

Hochschule für Musik Detmold (2007). *Ordnung für die Feststellung der besonderen Eignung in den Studiengängen Musik mit dem Abschluss Erste Staatsprüfung für das Lehramt an Gymnasien und Gesamtschulen*. Verfügbar unter: <http://www.hfmdetmold.de/de/dateien/docs/116/Eignungspruefungsordnung%20Schulmusik.pdf> [12.09.12]

Hochschule für Musik Detmold (o. J., a). *Mustertests – Musiktheorie*. Verfügbar unter: http://www.hfm-detmold.de/de/dateien/pdfs/mustertest_musiktheorie_sm.pdf [12.09.12]

Hochschule für Musik Detmold (o. J., b). *Mustertests – Hörfähigkeit*. Verfügbar unter: http://www.hfm-detmold.de/de/dateien/pdfs/mustertest_gehoerbildung_sm.pdf [12.09.12]

Hochschule für Musik, Theater und Medien Hannover (2010). *Ordnung über den Zugang und die Zulassung für Studiengänge mit dem Abschluss Bachelor of Music (B.Mus.) und für den Fächerübergreifenden Bachelorstudiengang Erstes Fach (Major) Musik an der Hochschule für Musik, Theater und Medien Hannover sowie für das Zweifach Musik im Rahmen des Studiengangs Sonderpädagogik an der Leibniz Universität Hannover (OZZ B.Mus.)*. Verfügbar unter: http://www.hmtmhannover.de/uploads/media/Zulassungsordnung_BMus_12.2010_10.pdf [12.09.12]

Hochschule für Musik, Theater und Medien Hannover (o. J.). *Musterklausur Bachelor und FüBA*. Verfügbar unter: http://www.hmtmhannover.de/uploads/media/Musterklausur_Bachelor_und_Fueba.pdf [12.09.12]

Holtmeier, L. (1997). Nicht Kunst? Nicht Wissenschaft? Zur Lage der Musiktheorie. *Musik & Ästhetik, 1* (1-2), 119-136.

Hornke, L. F. (2005). Die englische Fassung der DIN 33430. In K. Westhoff et al. (Hrsg.), *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430*. Lengerich: Pabst. Verfügbar unter: <http://www.bdp-verband.org/bdp/politik/clips/din33430en.pdf> [12.09.12]

Karpinski, G. S. (2000). *Aural skills acquisition. The development of listening, reading, and performing skills in college-level musicians*. Oxford: Oxford University Press.

Kersting, M. (2008). *Qualität in der Diagnostik und Personalauswahl – der DIN-Ansatz*. Göttingen: Hogrefe.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Ri-quarts, K., Tenorth, H.-E. & Vollmer, H. J. (Hrsg.) (2003). *Zur Entwicklung nationaler Bildungsstandards: Eine Expertise*. Berlin: BMBF.

Kühn, C. (2010). Musiktheorie lehren. Zu einer Umfrage an den deutschen Musikhochschulen. *Zeitschrift der Gesellschaft für Musiktheorie, 7*(1), 11-60. Verfügbar unter: <http://www.gmth.de/zeitschrift/artikel/505.aspx> [12.09.12]

Lehmann, A. C., & Hornberger, J. (2010). *Über den Zusammenhang von musikalischen Teilfertigkeiten im Studium und darüber hinaus*. Poster vorgestellt bei der Konferenz der Deutschen Gesellschaft für Musikpsychologie, Würzburg.

Mair, P., Hatzinger, R., & Maier, M. (2011). *Package 'eRm'*. Verfügbar unter: <http://cran.r-project.org/web/packages/eRm/eRm.pdf> [12.09.12]

Menke, J. (2010). Brauchen wir einen Kanon der Musiktheorie? *Zeitschrift der Gesellschaft*

- für *Musiktheorie*, 7(1), 61-70. Verfügbar unter: <http://www.gmth.de/zeitschrift/artikel/507.aspx> [12.09.12]
- Rektorenkonferenz der deutschen Musikhochschulen in der HRK (2009). *Leitbild der deutschen Musikhochschulen*. Verfügbar unter: <http://www.die-deutschen-musikhochschulen.de/content/media/dokumente/Leitbild%20Beschlussfassung.pdf> [12.09.12]
- Rolle, C. (2008). Musikalische Bildung durch Kompetenzerwerb? Überlegungen im Anschluss an den Entwurf eines Kompetenzmodells „Musik wahrnehmen und kontextualisieren“. *Zeitschrift für kritische Musikpädagogik, Sonderedition 2*, 42-59.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern [u. a.]: Verlag Hans Huber.
- Rost, J., & von Davier, M. (1994). A conditional Item-Fit Index for Rasch Models. *Applied Psychological Measurement*, 18, 171-182.
- Strobl, C. (2010). *Das Rasch-Modell. Eine verständliche Einführung für Studium und Praxis*. München und Mering: Rainer Hampp Verlag.
- Universität Bremen (2011). *Ordnung für die Aufnahmeprüfung im BA Musikpädagogik der Universität Bremen*. Verfügbar unter: http://www.musik.uni-bremen.de/fileadmin/mediapool/musik/dateien/Aufnahmepruefung/Aufnahmepruefungsordnung_MuPaed.pdf [12.09.12]
- Verband deutscher Musikschulen (1987). *Richtlinien für eine „Studienvorbereitende Ausbildung (SVA) als Abteilung an Musikschulen des VdM“*. Auf Anfrage per E-Mail (vdm@musikschulen.de) zu erhalten.
- Verband deutscher Musikschulen (1990). *Lehrplan Hörerziehung, Musiklehre und Musiktheorie*. Kassel: Gustav Bosse Verlag.
- Verband deutscher Musikschulen (2011). *Schülerzahl, Anzahl und Wochenstunden der Ensemble- und Ergänzungsfächer*. Verfügbar unter: <http://www.musikschulen.de/musikschulen/fakten/ensemble-undergaenzungsfaecher/index.html> [12.09.12]
- Weidner, V. (2010). Musiktheorie und Musikpädagogik. „Resonanzprobleme“ einer Beziehung. *Zeitschrift der Gesellschaft für Musiktheorie. Sonderausgabe*. Verfügbar unter: <http://www.gmth.de/zeitschrift/artikel/562.aspx> [12.09.12]

Autorinnen und Autoren:

Anna Wolf

Hochschule für Musik, Theater und Medien Hannover
Emmichplatz 1
30175 Hannover

Email: anna.wolf@hmtm-hannover.de

Reinhard Kopiez

Hochschule für Musik, Theater und Medien Hannover
Emmichplatz 1
30175 Hannover

Email: reinhard.kopiez@hmtm-hannover.de

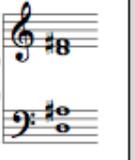
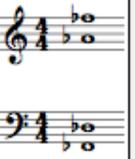
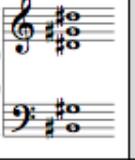
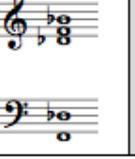
Friedrich Platz

Hochschule für Musik, Theater und Medien Hannover
Emmichplatz 1
30175 Hannover

Email: friedrich.platz@hmtm-hannover.de

Appendix A

Abb. 1: Lösungswahrscheinlichkeiten der geprüften Akkorde in Klausurversion 1 (K1) und Klausurversion 2 (K2)

	94%, kl. Dur-Sept- Akkord V2, Akk. 7		80%, Moll-Dreiklang V2, Akk. 4		73%, kl. Dur-Sept- Akkord V2, Akk. 14		65%, Moll-Dreiklang V2, Akk. 5		59%, halbv. Sept- Akkord V2, Akk. 8
	86%, kl. Moll-Sept- Akkord V2, Akk. 2		80%, überm. Dreiklang V2, Akk. 12		69%, kl. Dur-Sept- Akkord V2, Akk. 3		64%, verm. Dreiklang V1, Akk. 12		56%, gr. Dur-Sept- Akkord V1, Akk. 9
	86%, Moll-Dreiklang V2, Akk. 9		78%, Moll-Dreiklang V1, Akk. 1		69%, kl. Dur-Sept- Akkord V2, Akk. 10		63%, kl. Moll-Sept- Akkord V1, Akk. 5		52%, Moll-Dreiklang V1, Akk. 14
	84%, gr. Dur-Sept- Akkord V2, Akk. 6		74%, kl. Dur-Sept- Akkord V1, Akk. 7		68%, gr. Moll-Sept- Akkord V1, Akk. 15		63%, gr. Dur-Sept- Akkord V1, Akk. 13		45%, vollv. Sept- Akkord V1, Akk. 11
	82%, kl. Moll-Sept- Akkord V1, Akk. 10		73%, Dur-Dreiklang V2, Akk. 1		67%, Moll-Dreiklang V1, Akk. 2		62%, halbv. Sept- Akkord V1, Akk. 3		41%, halbv. Sept- Akkord V2, Akk. 15
	81%, kl. Dur-Sept- Akkord V1, Akk. 4		73%, Moll-Dreiklang V2, Akk. 13		66%, halb. Sept- Akkord V1, Akk. 8		61%, verm. Dreiklang V2, Akk. 11		41%, vollv. Sept- Akkord V1, Akk. 6

Appendix B

Abb. 1: Aggregierte Verteilung der Punktzahlen in den schriftlichen Zulassungsprüfungen Musiktheorie von 2009 bis 2011 (n = 302).

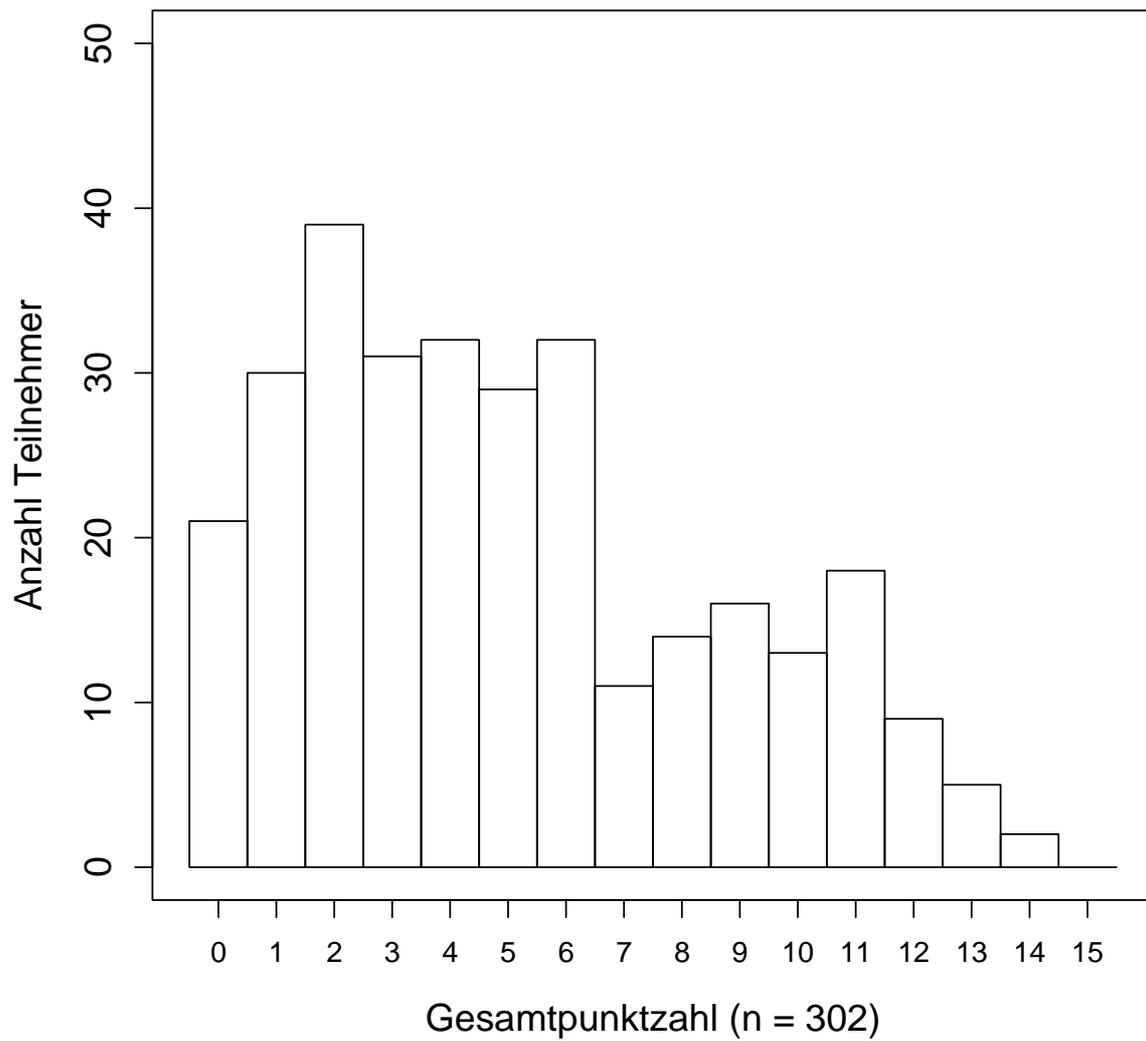


Abb. 2: Verteilung der Punktzahlen in der schriftlichen Zulassungsprüfung Musiktheorie des Jahres 2009 (n = 84). Der Schwellenwert für das Bestehen dieses Prüfungsteils liegt bei ≥ 7 Punkten.

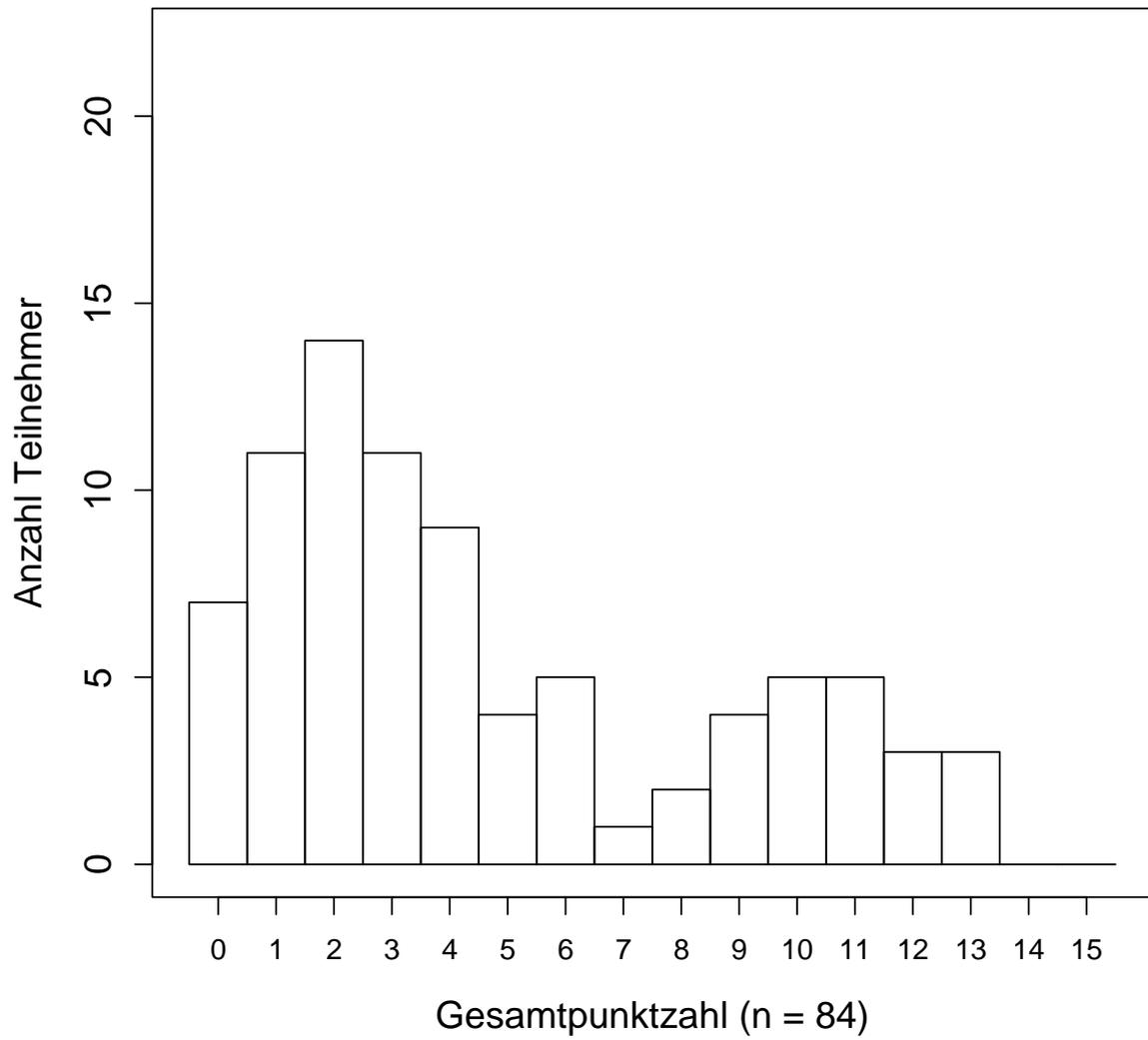


Abb. 3: Verteilung der Punktzahlen in der schriftlichen Zulassungsprüfung Musiktheorie des Jahres 2010 (n = 86). Der Schwellenwert für das Bestehen dieses Prüfungsteils liegt bei ≥ 7 Punkten.

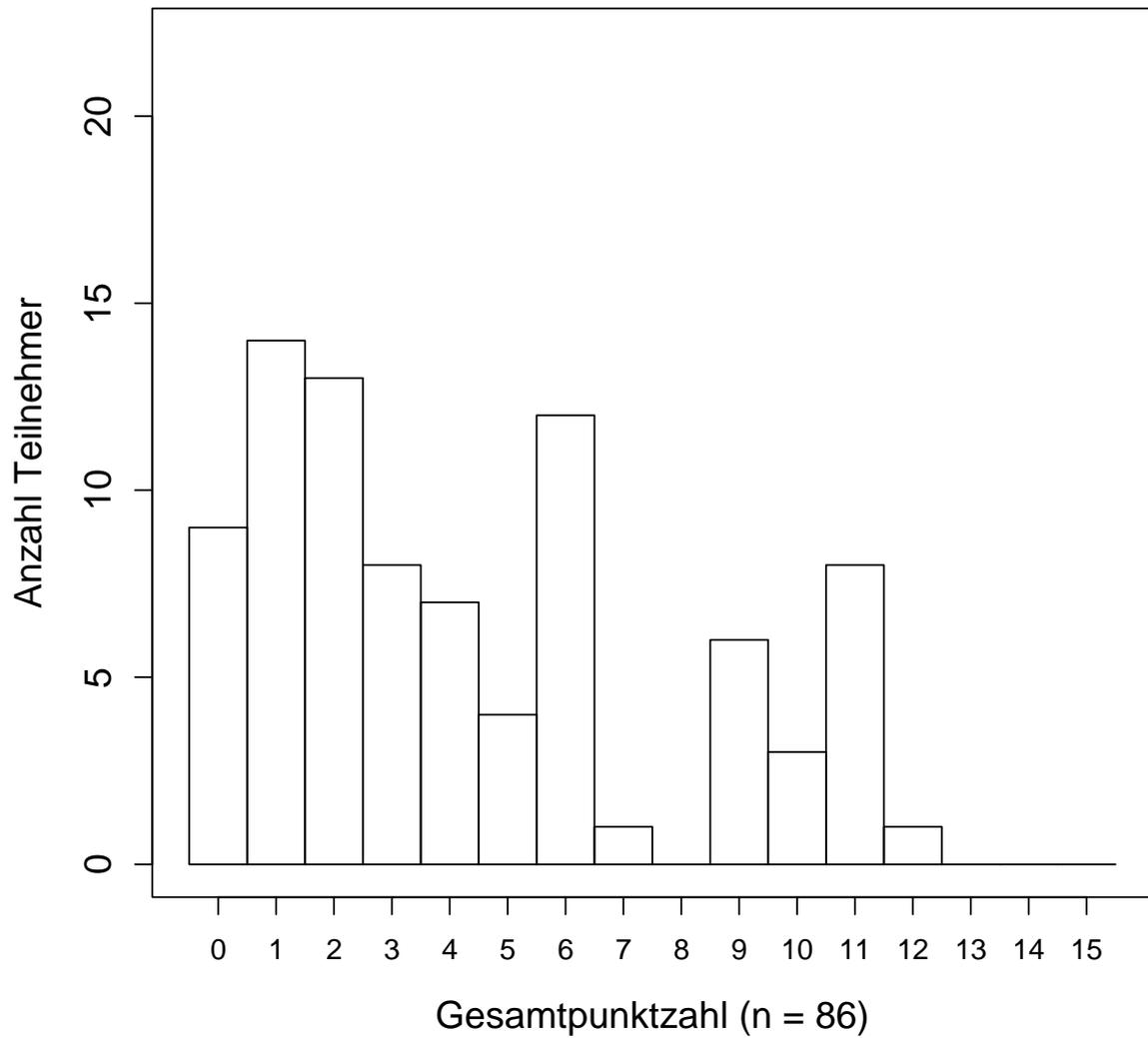
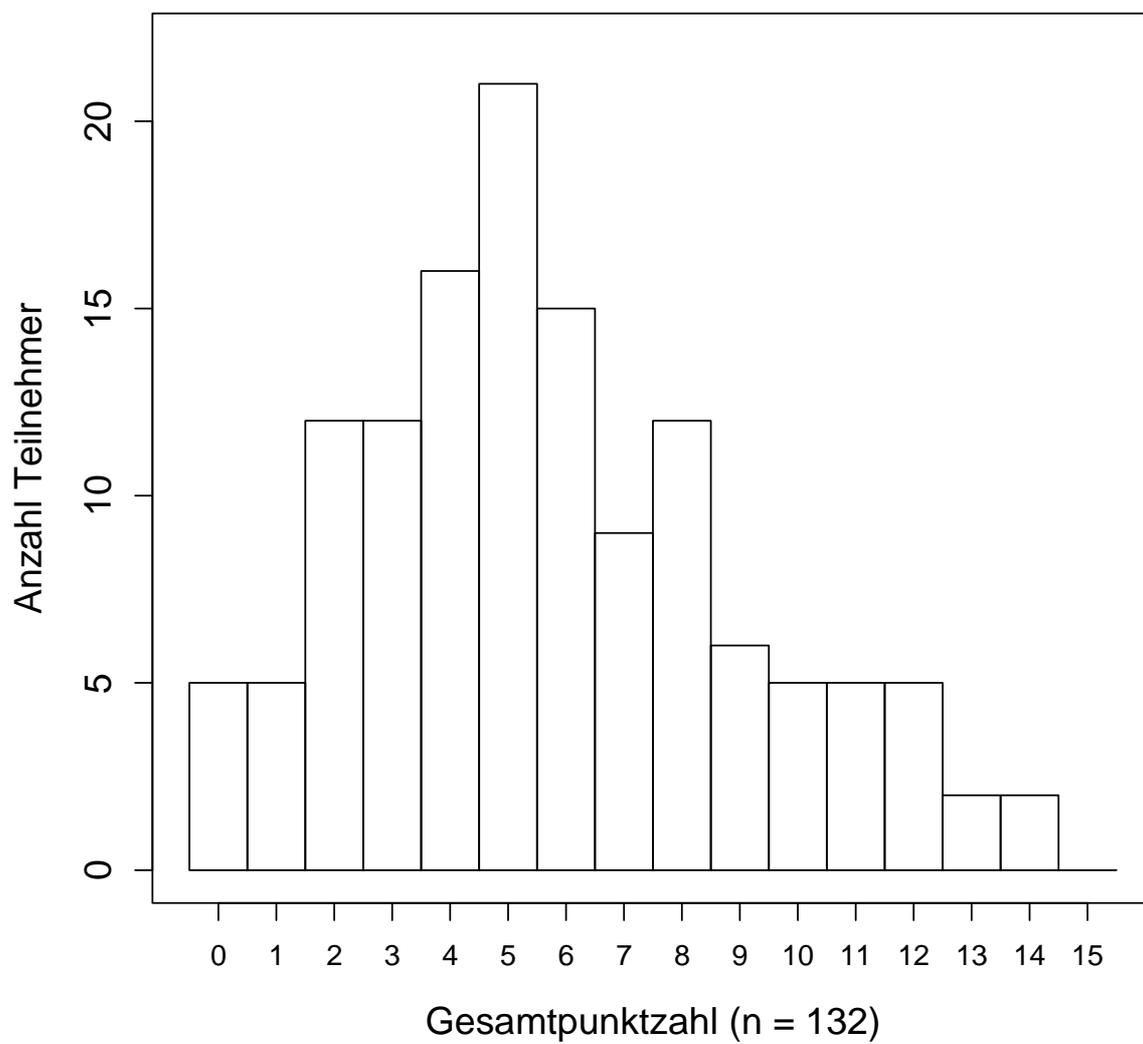


Abb. 4: Verteilung der Punktzahlen in der schriftlichen Zulassungsprüfung Musiktheorie des Jahres 2011. Der Schwellenwert für das Bestehen dieses Prüfungsteils liegt bei ≥ 7 Punkten.



Appendix C

Appendix C1: Itemkennwerte zur Klausurversion 1 (*kursiv*: nicht Rasch-konform)

Akkordnummer	Schwierigkeit (Logit)	Unteres CI (95 %)	Oberes CI (95 %)	Chi-Quadrat	<i>p</i>	MNSQ Infit (<i>t</i>)	MNSQ Outfit (<i>t</i>)
<i>Akk. 1</i>	-0,90	-0,27	-1,54	117,8	0,00	2,48	2,16
Akk. 2	-0,12	-0,68	0,44	57,3	0,82	-0,91	-0,51
<i>Akk. 3</i>	0,21	-0,33	0,75	106,2	0,00	3,11	2,82
Akk. 4	-1,13	-1,80	-0,47	46,2	0,98	-1,08	-0,01
Akk. 5	0,13	-0,41	0,68	68,0	0,48	-0,01	0,19
Akk. 6	1,33	0,81	1,86	76,0	0,24	0,63	0,20
Akk. 7	-0,58	-1,18	0,02	52,0	0,93	-1,08	-1,27
Akk. 8	-0,03	-0,59	0,52	45,2	0,99	-2,14	-1,77
<i>Akk. 9</i>	0,52	0,00	1,05	70,3	0,40	0,20	0,35
Akk. 10	-1,27	-1,95	-0,58	71,3	0,37	0,26	0,50
Akk. 11	1,11	0,59	1,63	61,9	0,68	-0,53	0,06
Akk. 12	0,05	-0,50	0,6	51,5	0,93	-1,55	-0,76
Akk. 13	0,13	-0,41	0,68	61,8	0,69	-0,58	-1,45
Akk. 14	0,75	0,23	1,27	59,7	0,76	-0,83	-0,85
Akk. 15	-0,21	-0,77	0,36	52,1	0,92	-1,33	-1,11

Appendix C2: Itemkennwerte zur Klausurversion 2 (*kursiv*: nicht Rasch-konform)

Akkordnummer	Schwierigkeit (Logit)	Unteres CI (95 %)	Oberes CI (95 %)	Chi-Quadrat	<i>p</i>	MNSQ Infit (<i>t</i>)	MNSQ Outfit (<i>t</i>)
<i>Akk. 1</i>	<i>0,18</i>	<i>0,88</i>	<i>0,54</i>	<i>64,1</i>	<i>0,02</i>	<i>2,05</i>	<i>1,54</i>
Akk. 2	-1,04	-1,94	-0,14	52,9	0,14	0,69	0,87
Akk. 3	0,44	-0,25	1,12	27,2	0,97	-2,10	-2,55
Akk. 4	-0,44	-1,24	0,35	23,3	0,99	-1,75	-1,22
<i>Akk. 5</i>	<i>0,68</i>	<i>0,02</i>	<i>1,35</i>	<i>53,9</i>	<i>0,12</i>	<i>1,23</i>	<i>0,97</i>
Akk. 6	-0,82	-1,68	0,04	43,4	0,45	0,14	0,04
Akk. 7	-2,24	-3,48	-1,01	40,1	0,60	0,25	0,39
Akk. 8	1,03	0,38	1,68	50,6	0,20	0,86	1,30
Akk. 9	-1,04	-1,94	-0,14	35,9	0,77	-0,29	-0,51
Akk. 10	0,44	-0,25	1,12	28,4	0,96	-1,91	-1,87
Akk. 11	0,91	0,26	1,57	46,4	0,33	0,38	1,00
Akk. 12	-0,44	-1,24	0,35	63,8	0,02	1,57	0,37
Akk. 13	0,18	-0,54	0,88	36,6	0,74	-0,71	-0,35
Akk. 14	0,18	-0,54	0,88	28,4	0,96	-1,74	-1,58
Akk. 15	2,01	1,34	2,68	38,7	0,66	-0,29	-0,39