

Hasselhorn, Johannes; Lehmann, Andreas C.

Leistungsheterogenität im Musikunterricht. Eine empirische Untersuchung zu Leistungsunterschieden im Bereich der Musikpraxis in Jahrgangsstufe 9

Niessen, Anne [Hrsg.]; Knigge, Jens [Hrsg.]: *Theoretische Rahmung und Theoriebildung in der musikpädagogischen Forschung*. Münster ; New York : Waxmann 2015, S. 163-176. - (Musikpädagogische Forschung; 36)



Quellenangabe/ Reference:

Hasselhorn, Johannes; Lehmann, Andreas C.: Leistungsheterogenität im Musikunterricht. Eine empirische Untersuchung zu Leistungsunterschieden im Bereich der Musikpraxis in Jahrgangsstufe 9 - In: Niessen, Anne [Hrsg.]; Knigge, Jens [Hrsg.]: *Theoretische Rahmung und Theoriebildung in der musikpädagogischen Forschung*. Münster ; New York : Waxmann 2015, S. 163-176 - URN: urn:nbn:de:0111-pedocs-126204 - DOI: 10.25656/01:12620

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-126204>

<https://doi.org/10.25656/01:12620>

in Kooperation mit / in cooperation with:



<http://www.ampf.info>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Anne Niessen, Jens Knigge (Hrsg.)

THEORETISCHE RAHMUNG UND THEORIEBILDUNG IN DER MUSIKPÄDAGOGISCHEN FORSCHUNG

THEORY FRAMEWORK AND DEVELOPMENT IN MUSIC EDUCATION RESEARCH

Musikpädagogische Forschung

Research in Music Education

Herausgegeben vom Arbeitskreis
Musikpädagogische Forschung e. V. (AMPF)

Band 36

Proceedings of the 36th Annual Conference of the
German Association for Research in Music Education

Anne Niessen, Jens Knigge (Hrsg.)

Theoretische Rahmung
und Theoriebildung in der
musikpädagogischen Forschung

Theory Framework and
Development
in Music Education Research



Waxmann 2015
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Musikpädagogische Forschung, Band 36 Research in Music Education, vol. 36

ISSN 0937-3993

ISBN 978-3-8309-3313-7

© Waxmann Verlag GmbH, Münster 2015
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Anne Breitenbach, Münster
Satz: Sven Solterbeck, Münster
Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des
Verlages in irgendeiner Form reproduziert oder unter Verwendung
elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

<i>Anne Niessen & Jens Knigge</i>	
Vorbemerkung	9
<i>Editors' note</i>	
<i>Matthias Proske</i>	
Unterricht als kommunikative Ordnung	
Eine kontingenzgewärtige Beschreibung	15
<i>Teaching as communication. A contingency-based approach</i>	
<i>Hermann J. Kaiser</i>	
Paradigma versus Denkstil	
Modi systematischer Historiographie in der Musikpädagogik	33
<i>Dimensions of historiography – Paradigma versus thought-style</i>	
<i>Thomas Busch</i>	
Das Wer, Wie und Was von (An-)Ordnungen	
Überlegungen zu Raumtheorie und Gerechtigkeit im Feld	
der Musikpädagogik	51
<i>Who arranges what – and how? Considerations regarding the theory</i>	
<i>of space and justice within the area of music education</i>	
<i>Verena Weidner</i>	
Soziologische Systemtheorie und Musikpädagogik	
Annäherungen an ein Theorieverhältnis	67
<i>Sociological Systems Theory and music education. Approaches to</i>	
<i>a relation between the theories</i>	

Adrian Niegot

Geltung – Gebrauch – Gehalt

Geltungstheoretische Überlegungen zu musikbezogenen historischen Sinnbildungen in musikpädagogischer Perspektive 81

Validity – usage – content. Reflections on validity regarding the construction of historical context from a music educational perspective

Malte Sachsse

Mensch und Musik im ‚Rahmen‘ der Theorien

Pluralisierung, Reflexion und Kritik als Aufgaben einer bildanthropologisch orientierten musikpädagogischen Forschung 95

Man and music in the ‘frame’ of theories. Pluralization, reflection and criticism as functions of a music pedagogical research orientated towards image-anthropology

Samuel Campos

Subjekte der Praxis – Praxis der Subjekte

Subjektivierung als Perspektive musikpädagogischer Unterrichtsforschung ... 111

Subjects of practice – practice of subjects. Subjectivation as a perspective for studies in music education

Monica Esslin-Peard, Tony Shorrocks & Graham F. Welch

The Art of Practice

Understanding the process of musical maturation through reflection 125

Lisa Knörzer, Christian Rolle, Robin Stark & Babette Park

„... er übertreibt und das macht mir seine Version zu nervös“

Einzelfallanalysen musikbezogener Argumentationen 147

“... he overacts and this makes his version too agitated for me”

A single-case analysis of music-related argumentations

Johannes Hasselhorn & Andreas C. Lehmann

Leistungsheterogenität im Musikunterricht

Eine empirische Untersuchung zu Leistungsunterschieden im Bereich der Musikpraxis in Jahrgangsstufe 9 163

Heterogeneity of achievements in the music classroom –

An empirical investigation regarding differences in practical musical competences in 9th grade

Melina Carmichael & Christian Harnischmacher

Ich weiß, was ich kann!

Eine empirische Studie zum Einfluss des musikbezogenen
Kompetenzerlebens und der Motivation von Schülerinnen und Schülern
auf deren Einstellung zum Musikunterricht 177

*I know my own skills. An empirical study regarding the influence
of music-related competence and motivation of students in their attitudes
towards music education at schools of general education*

Daniel Fiedler & Daniel Müllensiefen

Validierung des Gold-MSI-Fragebogens zur Messung
Musikalischer Erfahrungheit von Schülerinnen und Schülern
der Sekundarstufen an allgemeinbildenden Schulen 199

*Validation of the Gold-MSI questionnaire to measure musical sophistication
of German students at secondary education schools*

Michal Bischoff, Tim Sandkämper & Christoph Louven

Jugendliche und ‚Klassische Musik‘

Vorurteile und Klischees 221

Teenagers and classical music – Prejudices and clichés

Michael Ahlers & Andreas Seifert

Sprachliche Heterogenität im Musikunterricht 235

Language diversity in music lessons

Johannes Hasselhorn & Andreas C. Lehmann

Leistungsheterogenität im Musikunterricht

Eine empirische Untersuchung zu Leistungsunterschieden im Bereich der Musikpraxis in Jahrgangsstufe 9

Heterogeneity of achievements in the music classroom – An empirical investigation regarding differences in practical musical competences in 9th grade

Music teachers have often noticed – and even complained about – the heterogeneity of achievements in their general music classrooms. The purpose of this study in music competences was to quantify this range for the first time and attempt to explain individual differences in achievements in the area of practical music making. 420 German secondary school students (median age 15; 44 % male; 46 % learning to play a musical instrument outside of school) took a newly developed and validated competency test. The test was computer-based and administered in groups of up to 30 students. The three dimensions of music competence tested were singing, elementary music playing (on a specially designed application for tablet PC), and rhythm production. All music productions were recorded during testing and were later rated by two experts. Averaged ratings were then submitted to IRT analyses to arrive at person characteristics (WLE – weighted likelihood estimators).

Results showed that the spread in competence on all three dimensions was markedly larger than expected from normal distribution ($> +/- 2$ SDs). Gender and type of school attended generated a bias in favor of female students ($d = .41-1.01$), and students at a high school with a music emphasis (Musikgymnasium) ($d > .80$). Also, students who had private music lessons scored higher than those who didn't. Here, there was a clear effect of duration of instruction (2, 4, 6 or 8 years). The effect sizes ranged from .16 (singing, after 2 years) to 1.68 (elementary music making, after 8 years).

In summary, heterogeneity of musical competences is indeed very large in the music classroom: one reason is likely extracurricular music instruction of some students. It is unclear if the same heterogeneity can also be observed in other subject areas.

Die Arbeit mit Schülerensembles wie Schulchören, -orchestern oder -bands ist für viele Musiklehrer¹ der Höhepunkt der Arbeitswoche. Hier kann frei von curricularen Vorgaben intensiv musikalisch gearbeitet werden. Die aus dieser Arbeit häufig resultierenden Konzerte, bei denen die Schüler ihre Ergebnisse einem breiten Publikum vorstellen können, sind ein Gewinn für alle Beteiligten. Gerade in Instrumentalensembles sind die von den Schülern präsentierten Ergebnisse meist nur deshalb möglich, weil die mitwirkenden Schüler notwendige instrumentale Fähigkeiten und Fertigkeiten im privaten Instrumentalunterricht aufgebaut haben.

Schüler, die unter erheblichem Aufwand ihrer Zeit und dem Geld ihrer Eltern eine solche zusätzliche musikalische Ausbildung erhalten haben, sind von Musiklehrern nicht nur in den Schulensembles gern gesehen, sondern sie gelten aufgrund ihres im Instrumentalunterricht erworbenen bzw. vertieften Fachwissens vermutlich auch häufig im regulären Musikunterricht als Leistungsspitze bzw. als ‚Zugpferde‘ in ihren Klassen. Allerdings ist durch die Tatsache, dass sich Schüler in ihrer Freizeit ganz unterschiedlich in Art und Intensität mit Musik beschäftigen oder auseinandersetzen, im Musikunterricht eine Situation vorzufinden, in der Musiklehrer mit Schülern stark unterschiedlicher Voraussetzungen und Fähigkeiten konfrontiert sind.

Diese Situation scheint zu einer besonders großen Heterogenität der Lerngruppen zu führen, auch im Vergleich zu anderen Fächern, die seit Jahrzehnten immer wieder berichtet wird (vgl. Hoene & Thurmann, 2011). Da zu diesem Sachverhalt bisher kaum quantitativ empirische Daten vorliegen, erscheint es lohnenswert, den Versuch zu unternehmen, Leistungsheterogenität im Musikunterricht zu quantifizieren.²

Hintergrund

Leistungsheterogenität scheint hierzulande von Lehrern grundsätzlich als etwas Negatives wahrgenommen zu werden, das die Arbeit erschwert (vgl. Budde, 2012; Tillmann, 2007; Reh, 2005). Ähnliches kann vermutlich für weite Teile der Musikpädagogik angenommen werden. Vogt (2012a) identifiziert zumindest in musikpädagogischen Konzeptionen wie der Auditiven Wahrnehmungserziehung oder dem Aufbauenden Musikunterricht „Homogenisierungstendenzen“, die letztlich auf einen homogenen Lernbeginn abzielen. Es sei dahingestellt, ob solche Tendenzen Wunsch oder Absicht von Entwicklern der genannten Konzeptionen gewesen sind oder nicht.

-
- 1 Aus Gründen der besseren Lesbarkeit wird in diesem Beitrag im vollen Bewusstsein der Existenz zweier gleichberechtigter Geschlechter in der Regel die zum Einschluss beider Geschlechter vorgesehene grammatisch korrekte Form des generischen Maskulinums verwendet.
 - 2 Die dieser Arbeit zugrunde liegenden Daten entstammen dem von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt KOPRA-M (LE 2204/6–1) des Zweitautors.

Doch auch wenn der Lernbeginn ein homogener sein kann, wäre es unrealistisch anzunehmen, dass sich Lerngruppen (die im Übrigen im Laufe der Schulzeit mehrfach neu zusammengestellt werden) im Musikunterricht über 9 bis 13 Schuljahre hinweg homogen im Lernverhalten und -erfolg darstellen werden. Verschiedene Veranlassungen, Interessen und Rahmenbedingungen führen zwangsläufig zu Leistungsdivergenzen, die dann in der Gruppe als Heterogenität wahrgenommen werden.

Als Hauptgrund für Leistungsheterogenität im Musikunterricht wird häufig angeführt, dass sich einige Schüler außerhalb des regulären Musikunterrichts in Form von außerschulischem Instrumentalunterricht intensiv und praktisch mit Musik auseinandersetzen und andere nicht (Hoene & Thurmann, 2011). Aktuell spielen etwa 36 % der 13- bis 18-Jährigen in Deutschland ein Musikinstrument (Autorengruppe Bildungsberichterstattung, 2012), für die das genannte Argument Anwendung finden könnte. Die schulischen Leistungsvorteile dieser Schüler konnten auch bereits für Bereiche der Musikwahrnehmung empirisch belegt werden (vgl. Jordan, 2014; Persky, Sandene & Askew, 1997).

Es ist zu vermuten, dass das Erlernen eines Musikinstruments vor allem Einfluss auf den musikpraktischen Bereich des regulären Musikunterrichts hat und dass Instrumentalschüler musikpraktische Aufgaben erfolgreicher lösen als ihre Mitschüler. Bisher lagen zu dieser Vermutung, die sich auf seit Jahren wiederkehrende Berichte von Musiklehrern stützt (vgl. Hoene & Thurmann, 2011), keine empirischen Daten vor. Ein Grund dafür war vor allem auch, dass bisher kein geeignetes Instrument zur Messung musikpraktischer Leistungen in größeren Stichproben vorlag. Im Rahmen der Kompetenzorientierung, die – wenn auch zögerlich – auch die Musikpädagogik erfasst hat (vgl. Knigge, 2014), ist der Bereich der Musikpraxis schon früh als potentieller Kompetenzbereich identifiziert worden (Niessen, Lehmann-Wermser, Knigge & Lehmann, 2008). Ein auf Basis von Literaturrecherche erstelltes Strukturmodell musikpraktischer Kompetenzen, das aus den drei Dimensionen *Gesang*, *instrumentales Musizieren* und *Rhythmusproduktion* besteht (Hasselhorn & Lehmann, 2014), konnte empirisch validiert werden, sodass ein Gruppentest zur Bestimmung von Kompetenzwerten auf diesen drei Dimensionen für Schüler der Jahrgangsstufe 9 vorliegt (Hasselhorn, 2014). Dieser Test misst auf allen drei Dimensionen mit sehr hoher Zuverlässigkeit (alle EAP/PV-Reliabilitäten > .90).

Dieses Testverfahren erlaubt die empirische Untersuchung der Fragestellungen, ob (a) für den Bereich der Musikpraxis in Jahrgangsstufe 9 tatsächlich eine statistisch überdurchschnittliche Leistungsheterogenität vorliegt und inwiefern (b) der außerschulische Instrumentalunterricht für diese Leistungsheterogenität verantwortlich gemacht werden kann. Im Zuge der zweiten Fragestellung sollen außerdem vergleichend das Geschlecht und die Schulform Berücksichtigung finden, die im Kontext schulischer Leistungen ebenfalls als Gründe für Heterogenität angeführt werden (z. B. Gröhlich, Scharenberg & Bos, 2009).³

3 Natürlich gibt es weitere potentielle Einflussgrößen auf musikpraktische Leistungsheterogenität wie elterliche Unterstützung, sozioökonomischer Status, Migrationshinter-

Methode

Zur Beantwortung der gestellten Fragen wurden Daten aus der Erhebung verwendet, die zur Validierung des Strukturmodells musikpraktischer Kompetenzen durchgeführt worden war (Hasselhorn, 2014).

Stichprobe

Teilnehmer an der Studie waren 420 Schüler von sechs verschiedenen Schulen aus der Umgebung Würzburgs (zwei musische Gymnasien, drei Gymnasien, eine Realschule). Zum Zeitpunkt der Erhebung besuchten sie die 9. Klasse (Median_{Alter} = 15 Jahre). 43.8 % der Probanden waren männlich und 46.2 % erhielten Instrumentalunterricht.

Durchführung

Im Vorfeld der Testungen füllten die Schüler einen Begleitfragebogen aus, in dem unter anderem Geschlecht, Schulform und Teilnahme an außerschulischem Instrumentalunterricht angegeben werden sollten. Die ausgefüllten Fragebögen sowie die ausgefüllten Einverständniserklärungen der Eltern brachten die Schüler zum Test mit. Der Test selbst wurde im Klassenverband in den teilnehmenden Schulen durchgeführt und dauerte etwa 90 Minuten. Die Schüler betraten den vorbereiteten Testraum, in dem jeder einen individuellen Arbeitsplatz zugewiesen bekam. Diese Arbeitsplätze bestanden aus einem Laptop, über den die Aufgaben präsentiert wurden, und einem Headset, über das sowohl Aufgabenstellungen und Tonbeispiele abgespielt als auch der Gesang der Schüler aufgenommen wurde. Mittels eines Tablet-PC, auf dem eine speziell für diesen Test entwickelte Applikation, die CMG-App (Hasselhorn & Grollmisch, 2014), installiert war, konnten die Schüler die Instrumental- und Rhythmusaufgaben bearbeiten. Der Tablet-PC war mit einer speziellen Halterung auf den Laptop aufgesetzt, um für eine größtmögliche Standardisierung des Versuchsaufbaus zu sorgen. Der gesamte Arbeitsplatz war von einer hölzernen Trennwand umgeben, sodass sich die Schüler untereinander nicht sehen konnten. Alle Schülerlaptops waren in einem Netzwerk mit einem Zentralrechner verbunden, über den der Testleiter zum einen den technischen Ablauf kontrollieren und zum anderen Aufgaben zentral zuteilen und starten konnte (für Details vgl. Hasselhorn, 2014).

Zu jeder der drei Dimensionen musikpraktischer Kompetenzen gab es eine Reihe standardisierter, gemeinsam mit Musiklehrern erarbeiteter Aufgaben, die

grund etc. Aufgrund forschungsökonomischer und datenschutzrechtlicher Beschränkungen im Zuge der Testvalidierung, deren Daten in diesem Beitrag analysiert werden, konnten diese und weitere Variablen jedoch keine Berücksichtigung finden.

in der Reihenfolge *instrumentales Musizieren*, *Rhythmusproduktion* und *Gesang* durchgeführt wurden. Den jeweiligen Blöcken war jeweils eine kurze Einführung samt Probeaufgaben vorangestellt, um den Versuchsablauf einzuüben, anschließend begannen die Testaufgaben. Innerhalb einer Aufgabe hörten die Schüler über die Kopfhörer zunächst die vorgelesene Aufgabenstellung, anschließend wurde bei einigen Aufgaben ein Hörbeispiel abgespielt. Es folgten einige Sekunden der Stille, in der sich die Schüler individuell auf die Bearbeitung vorbereiten konnten. Diese Stille wurde mit einem Startsignal beendet, das als zeitliche Markierung für den Beginn der eigentlichen Testaufgabe fungierte. Zu jeder Aufgabe wurde dann ein Playback bzw. Playalong abgespielt, zu dem die Schüler je nach Aufgabe singen oder spielen sollten. Diese Playbacks oder Playalongs waren nötig, um die Aufgaben in zeitlicher Hinsicht zu standardisieren. Eventuelle Notenbeispiele, die zu den Aufgaben gehörten, wurden auf dem Bildschirm des Laptops abgebildet. Die Schülerantworten, also das, was die Schüler sangen oder spielten, wurden aufgezeichnet und auf dem Zentralrechner gespeichert.

Datenaufbereitung

Die Schüleraufnahmen wurden zu einem späteren Zeitpunkt von Experten auf 6-stufigen Skalen bewertet, wobei jede Aufnahme von mindestens zwei Ratern beurteilt wurde. Mit Hilfe von Intraklassenkorrelationen (vgl. Wirtz & Caspar, 2002) wurde kontrolliert, dass die Raterurteile ausreichend übereinstimmten. Da dies für alle Aufgaben der Fall war, wurden anschließend für jede Aufnahme die Mittelwerte der Raterurteile gerundet, sodass für jeden Schüler für jede bearbeitete Aufgabe eine ganzzahlige Bewertung vorlag.

Diese Daten wurden mit Methoden der probabilistischen Testtheorie (auch „item response theory“: IRT) weiter aufbereitet. Hierbei wird in einem mathematischen Modell – in diesem Fall dem ordinalen Raschmodell – die individuelle Personenfähigkeit zur Aufgabenschwierigkeit in Beziehung gesetzt. Dadurch kann berücksichtigt werden, dass auch leistungsstarke Schüler gelegentlich eine leichte Aufgabe falsch bzw. vergleichsweise schlecht lösen oder dass umgekehrt leistungsschwache Schüler in Einzelfällen schwere Aufgaben richtig bzw. vergleichsweise gut lösen. Auf Grundlage der Antworten können spezielle Statistikprogramme wie das in dieser Studie verwendete *Conquestz* dann die wahrscheinlichsten Kompetenzausprägungen für alle Dimensionen schätzen (vgl. Bühner, 2011, Kap. 7). Diese Kompetenzschätzer (sogenannte WLEs) – ein Wert für jeden Schüler und jede Dimension musikpraktischer Kompetenzen – wurden als unabhängige Variable in den statistischen Analysen zur Klärung der hier behandelten Forschungsfragen herangezogen.

Ergebnisse

Um der Frage nach der Leistungsheterogenität hinsichtlich der Musikpraxis nachzugehen, wurden zunächst in einem rein deskriptiven Schritt die Verteilungen der Kompetenzschätzer (WLEs) getrennt für alle drei Dimensionen mit der Standardnormalverteilung verglichen. Die Standardnormalverteilung kann, bei ausreichend großer Stichprobe, als Vergleich herangezogen werden, um zu entscheiden, ob eine Verteilung überdurchschnittlich breit ist oder nicht. Den Normwerten entspricht demnach eine Verteilung mit einer Breite von insgesamt sechs Standardabweichungen (Bühner & Ziegler, 2009, S. 56). Die WLEs werden im Rahmen der IRT-Verfahren auf der sogenannten Logit-Skala abgetragen. Diese Skala hat den Mittelwert 0 und analog der Normalverteilung erwartbare Werte zwischen -3 und +3 (Moosbrugger, 2008, S. 254). Die Logit-Einheiten können also direkt als Standardabweichungen interpretiert werden.

Die aus den Daten abzulesenden Spannweiten der WLE-Verteilungen (s. Tab. 1) sind allerdings als problematisch anzusehen. Da Testverfahren grundsätzlich dazu neigen in der Nähe des Mittelpunktes sehr gut und an den Extrema deutlich schlechter zu differenzieren, können einzelne Ausreißerwerte das Ergebnis derart beeinflussen, dass die wahre Spannweite überschätzt wird (vgl. Reinders, Ditton, Gräsel & Gniewosz, 2011, S. 75). Um dieses Problem zu umgehen, wurden zusätzlich in Tab. 1 die Spannweiten der mittleren 95 % der WLE-Werte angegeben. Dies wird durch die Eliminierung der oberen und unteren 2,5 % der Werte erreicht. Der übrig bleibende Datenpool sollte nun keine Ausreißer mehr aufweisen. Ein weiterer Vorteil ist, dass die mittleren 95 % einer Standardnormalverteilung ungefähr den mittleren 4 Standardabweichungen um den Mittelpunkt entsprechen. Bezogen auf Intelligenz (Mittelwert: $M_{IQ} = 100$, Standardabweichung $SD_{IQ} = 15$) entspräche das einem Bereich von 70 bis 130 IQ-Punkten, in dem sich 95 % aller Personen befinden. Die restlichen 5 % haben jeweils zur Hälfte entweder einen IQ über 130 oder einen IQ unter 70. Die so gefundenen Spannweiten sollten also eine Breite von etwa vier Logiteinheiten aufweisen, um der Standardnormalverteilung zu entsprechen.

Tabelle 1: Spannweiten der empirisch gefundenen WLE-Verteilungen aller drei Dimensionen musikpraktischer Kompetenz des gesamten Datensatzes ($n = 420$) und die jeweils mittleren 95 % des Datensatzes ($n = 398$) in Logiteinheiten.

Dimension	100 %			95 %		
	Minimum	Maximum	Spannweite	Minimum	Maximum	Spannweite
<i>Gesang</i>	-4.90	5.98	10.88	-3.63	4.05	7.68
<i>Instrumentales Musizieren</i>	-3.97	5.04	9.01	-2.38	3.16	5.54
<i>Rhythmusproduktion</i>	-5.21	2.58	7.79	-2.43	1.94	4.37

Lediglich die Leistungsverteilung der Dimension *Rhythmusproduktion* ist annähernd standardnormalverteilt, wie der letzten Spalte von Tabelle 1 zu entnehmen ist. In Bezug auf Rhythmusleistungen kann in Jahrgangsstufe 9 also eine gewöhnliche Leistungsverteilung erwartet werden, da die gefundene Spannweite mit 4.37 ungefähr der zu erwartenden Spannweite von 4.00 entspricht. Im Gegensatz dazu zeigen sich die gefundenen Verteilungen auf den Dimensionen *Gesang* und *instrumentales Musizieren* deutlich breiter als die Standardnormalverteilung. Mit einer Breite von 5,54 Logiteinheiten ist die empirisch gefundene Leistungsstreuung auf der Dimension *instrumentales Musizieren* 38.5 % breiter, auf der Dimension *Gesang* mit 7.68 sogar 92 % breiter, als das zu erwarten gewesen wäre. Zur Anschaulichkeit übertragen auf Intelligenzwerte wären also nicht etwa die üblichen Werte zwischen 70 und 130 zu finden, sondern sogar Werte zwischen 60 und 140 (*instrumentales Musizieren*) bzw. Werte zwischen 40 und 160 (*Gesang*).

Eine generelle überdurchschnittliche Leistungsheterogenität für den Bereich der Musikpraxis konnte also für die Dimensionen *Gesang* und *instrumentales Musizieren* festgestellt werden, nicht aber für die Dimension der *Rhythmusproduktion*. Es sollte aber darüber hinaus der Frage nachgegangen werden, inwiefern die üblicherweise zur Erklärung solcher Leistungsdifferenzen herangezogenen Hintergrundvariablen *außerschulischer Instrumentalunterricht*, *Geschlecht* und *Schulform* für diese Leistungsdifferenzen verantwortlich sind. Zur Klärung dieser Fragestellung wurden zunächst Mittelwerte und Standardabweichungen für die jeweiligen Subgruppen dieser Variablen berechnet (s. Tab. 2).

Tabelle 2: Mittelwerte und Standardabweichungen (in Klammern) der WLEs für die drei Dimensionen *Gesang*, *instrumentales Musizieren* und *Rhythmusproduktion* getrennt nach *Geschlecht*, *Schulform* und *außerschulischem Instrumentalunterricht*.

Dimension	Geschlecht		Schulform			außerschulischer Instrumentalunterricht	
	m (n = 184)	w (n = 236)	musisches Gymn. (n = 132)	Gym- nasium (n = 248)	Real- schule (n = 40)	ja (n = 194)	nein (n = 226)
Gesang	-1.09 (1.92)	.84 (1.89)	1.35 (1.97)	-.44 (1.87)	-1.81 (1.70)	1.03 (1.99)	-.89 (1.83)
Instrumentales Musizieren	-.52 (1.34)	.43 (1.42)	.93 (1.49)	-.37 (1.28)	-.62 (1.02)	.81 (1.41)	-.69 (1.13)
Rhythmusproduktion	-.26 (1.17)	.18 (.98)	.57 (.92)	-.24 (1.04)	-.55 (1.12)	.58 (.82)	-.52 (1.83)

Die offenkundig vorhandenen Mittelwertsunterschiede wurden mittels *T*-Tests zunächst auf statistische Signifikanz überprüft, um sicherzustellen, dass es sich tatsächlich um überzufällige Unterschiede handelte. Für auf diese Weise identifizierte Unterschiede wurden zusätzlich *d*-Werte berechnet, um diese Effekte zu quantifizieren. Dabei gilt ein *d* ab .2 als kleiner, ab .5 als mittlerer und ab .8 als großer Effekt (Schäfer, 2011, S. 81).

Es zeigte sich, dass auf allen drei Dimensionen musikpraktischer Kompetenzen teils erhebliche Geschlechtsunterschiede vorlagen. Werte von $t(418) = 10.28$ ($p < .01$), $d = 1.01$ (*Gesang*), $t(418) = 6.92$ ($p < .01$), $d = .68$ (*instrumentales Musizieren*) und $t(418) = 4.19$ ($p < .01$), $d = .41$ (*Rhythmusproduktion*) verdeutlichen, dass in der Jahrgangsstufe 9 Mädchen gegenüber Jungen im Bereich des Singens im Rahmen eines großen Effektes, im Bereich des Instrumentalspiels im Rahmen eines mittleren und im Bereich des rhythmischen Musizierens im Rahmen eines kleinen Effektes überlegen sind.

Zur Berechnung des Einflusses der Schulform wurden für jede Dimension musikpraktischer Kompetenzen je zwei *T*-Tests berechnet. Zunächst wurden die Schüler musischer Gymnasien mit Schülern von Gymnasien verglichen, anschließend noch einmal die Schüler von Gymnasien mit Realschülern. Aufgrund der Alpha-Fehler-Kumulierung wurde für diese *T*-Tests das Signifikanzniveau pro Einzeltest auf .025 abgesenkt, um eine globale Fehlerwahrscheinlichkeit von 5 % zu wahren. Dabei konnte im Bereich Gesang für den Unterschied zwischen musischen Gymnasien und Gymnasien ein Wert von $d = .94$ ($t[378] = 8.71$; $p < .01$) und für den Unterschied zwischen Gymnasien und Realschulen ein Wert von $d = .75$ ($t[286] = 4.37$; $p < .01$) gefunden werden. Die untersuchten Schüler musischer Gymnasien zeigten demnach eine erheblich bessere Gesangsleistung als Schüler normaler Gymnasien, die wiederum im Rahmen eines mittleren Effektes besser sangen als die gleichaltrigen Realschüler. Für den Bereich des Instrumentalspiels konnte lediglich ein großer Effekt für den Unterschied zwischen Schülern musischer Gymnasien und Gymnasiasten von $d = .96$ ($t[378] = 8.91$; $p < .01$) gefunden werden. Zwischen den Schulformen Gymnasium und Realschule konnte hingegen keine statistisch signifikante Differenz festgestellt werden ($t[286] = 1.19$; $p = .23$). Im Instrumentalspiel zeigten also Schüler musischer Gymnasien erheblich bessere Leistungen als alle anderen Schüler, zwischen Schülern von Gymnasien und Realschulen konnte hingegen in unserer Stichprobe kein Unterschied nachgewiesen werden. Ähnliche Befunde zeigten sich für den rein rhythmischen Bereich. Hier zeigten die Schüler der musischen Gymnasien im Mittel stärkere Leistungen als die gleichaltrigen Gymnasiasten im Rahmen eines großen Effektes ($d = .81$; $t[378] = 7.48$; $p < .01$), zwischen Schülern von Gymnasium und Realschule war hingegen kein statistisch bedeutsamer Unterschied feststellbar ($t[286] = 1.73$; $p = .09$).

Ein deutlicher Einfluss des außerschulischen Instrumentalunterrichts konnte ebenfalls nachgewiesen werden. Im Bereich des Singens erzielten Schüler mit außerschulischem Instrumentalunterricht im Mittel deutlich bessere Ergebnisse als ihre Klassenkameraden in der Größe von $d = 1.01$ ($t[418] = 10.31$; $p < .01$). Im Bereich

des Instrumentalspiels fiel der Unterschied mit einem d von 1.17 ($t[418] = 11.97$; $p < .01$) ebenso wie im Bereich der Rhythmusproduktion mit einem d von 1.17 ($t[418] = 12.00$; $p < .01$) tendenziell noch etwas deutlicher aus. Diese Ergebnisse sprechen dafür, dass Schüler, die außerschulischen Instrumentalunterricht erhalten, in der Jahrgangsstufe 9 bei musikpraktischen Tätigkeiten im Mittel tatsächlich über eine Standardabweichung bessere Leistungen erzielen als Schüler ohne diese zusätzliche musikalische Beschäftigung.

Unberücksichtigt bei dieser Art der Auswertung bleibt allerdings die Intensität des außerschulischen Instrumentalunterrichts. Es ist anzunehmen, dass intensivere musikalische Beschäftigung, die zusätzlich zum regulären Musikunterricht betrieben wird, positivere Auswirkungen auf die für die Schule relevanten musikpraktischen Kompetenzen hat als weniger intensive Beschäftigung mit einem Musikinstrument. Die Intensität instrumentaler Ausbildung lässt sich allerdings bis heute kaum quantifizieren. Die einzige Annäherung an einen solchen Wert, die momentan unter vertretbarem Aufwand möglich ist, ist die Messung der Dauer des bisherigen Instrumentalunterrichts. Aus diesem Grund wurde für eine detailliertere Analyse die gesamte Stichprobe in fünf Subgruppen nach der bisherigen Dauer des Instrumentalunterrichts in Zwei-Jahres-Schritten (2, 4, 6, 8 Jahre) unterteilt. Eine genauere Unterteilung war aufgrund der dann zu geringen Fallzahlen für statistische Analysen nicht geeignet. Für diese Subgruppen wurden wiederum Mittelwerte und Standardabweichungen der WLEs für alle drei Dimensionen musikpraktischer Kompetenzen berechnet sowie d -Werte zur standardisierten Quantifizierung der Differenzen zwischen den Mittelwerten der vier Dauerngruppen von Schülern mit und ohne Instrumentalunterricht (s. Tab. 3).

Tabelle 3: Mittelwerte und Standardabweichungen der WLEs für die drei Dimensionen musikpraktischer Kompetenzen nach Dauer des außerschulischen Instrumentalunterrichts mit d -Werten zur Beschreibung des standardisierten Abstands zur Gruppe von Schülern ohne (kein) Instrumentalunterricht

Dimension	Dauer des außerschulischen Instrumentalunterrichts in Jahren									
	kein ($n = 226$)		ca. 2 ($n = 25$)		ca. 4 ($n = 50$)		ca. 6 ($n = 56$)		ca. 8 ($n = 54$)	
	M (SD)	M (SD)	d							
Gesang	-0.89 (1.83)	-0.59 (1.98)	.16	.75 (1.95)	.89	1.44 (1.98)	1.25	1.50 (1.72)	1.32	
Instrumentales Musizieren	-0.67 (1.13)	-0.40 (.86)	.24	.67 (1.32)	1.15	.82 (1.23)	1.30	1.34 (1.43)	1.68	
Rhythmusproduktion	-0.52 (1.03)	-0.06 (.67)	.46	.52 (.86)	1.04	.67 (.69)	1.23	.77 (.87)	1.29	

Die Ergebnisse zeigen, dass die erheblichen Leistungsdifferenzen in Klassenstufe 9 auf diejenigen Schüler zurückzuführen sind, die seit mindestens vier Jahren außerschulischen Instrumentalunterricht erhalten. Lediglich für die Dimension *Rhythmusproduktion* kann bereits für Schüler mit nur zwei Jahren Instrumentalunterricht eine bessere Leistung im Rahmen eines kleinen, aber deutlichen Effekts gefunden werden.

Abschließend wurde für jede Dimension eine Regressionsanalyse durchgeführt, um zu überprüfen, wie stark der Einfluss der Dauer des Instrumentalunterrichts auf die drei Dimensionen musikpraktischer Kompetenzen unter Konstanzhaltung von Geschlecht und Schulform war. Vorbereitend wurden die Interkorrelationen der drei Prädiktoren berechnet. Dabei ergaben sich statistisch signifikante Korrelationen zwischen den Variablen Geschlecht und Schulform von $r = .18$ sowie zwischen den Variablen Geschlecht und Dauer des Instrumentalunterrichts von $r = .17$. Diese Korrelationen zeigen, dass Mädchen tendenziell häufiger musische Gymnasien oder Gymnasien besuchen als Jungen und dass Mädchen in Klasse 9 tendenziell länger Instrumentalunterricht als ihre männlichen Klassenkameraden haben. Für die Korrelationen zwischen den Variablen Dauer des Instrumentalunterrichts und Schulform ergab sich ein statistisch signifikantes r von $.35$, was dahingehend interpretiert werden kann, dass Schüler am musischen Gymnasium in Klasse 9 am längsten ein Instrument erlernen und Realschüler am kürzesten. Darüber hinaus bedeuten die Korrelationen, dass etwaige β -Gewichte einer Regressionsanalyse nicht als Effektgrößen interpretiert werden können.

Da vor allem der alleinige Effekt der Dauer des Instrumentalunterrichts von Interesse war, wurden drei hierarchische Regressionen durchgeführt, bei denen in einem ersten Schritt die Prädiktoren Geschlecht und Schulform gemeinsam aufgenommen und in einem zweiten Schritt überprüft wurde, wie viel zusätzliche Varianz auf den Testwerten der musikpraktischen Kompetenzdimensionen von der Dauer des Instrumentalunterrichts aufgeklärt werden konnte. Für die Dimension *Gesang* ergab sich insgesamt ein Modell mit einem R^2 von $.45$ und einer alleinigen Aufklärungsquote für die Dauer des Instrumentalunterrichts von 8.9% . In der Dimension *instrumentales Musizieren* konnten insgesamt 39.0% Varianz aufgeklärt werden, von denen 16.7% ausschließlich durch die Dauer des Instrumentalunterrichts erklärt werden konnten. Bezüglich der *Rhythmusproduktion* war das gesamte R^2 mit $.30$ etwas niedriger, davon entfielen auf die Dauer des Instrumentalunterrichts allerdings 15.6% .

Diskussion

In diesem Beitrag wurden zwei Fragen untersucht. Die erste Frage war die nach der Existenz einer überdurchschnittlichen Leistungsheterogenität in der Musikpraxis in Jahrgangsstufe 9. Diese Frage kann für die musikpraktischen Kompetenzdimensionen *Gesang* und *instrumentales Musizieren* bejaht werden. Es konnten eindeuti-

ge Hinweise dafür gefunden werden, dass die Streuungsbreite der Leistungen bzw. Leistungsdispositionen (ausgedrückt als WLEs aus einem eigens entwickelten Kompetenztest) auf diesen beiden Dimensionen das zu erwartende Maß bei weitem übersteigt. Einschränkend muss festgehalten werden, dass das nicht für die rhythmische Dimension gilt, deren Streuungsbreite derjenigen der Standardnormalverteilung entspricht. Wenn man allerdings davon ausgeht, dass musikpraktische Arbeit in der Schule vor allem in den Bereichen Gesang und instrumentales Musizieren stattfindet, müssen Musiklehrer täglich mit den gefundenen extremen Heterogenitäten umgehen. Da dieser Zustand bereits seit vielen Jahren kontinuierlich berichtet wird, ist dies allerdings keine unbedingt neue Erkenntnis. Vielmehr handelt es sich erstmalig um eine empirische Quantifizierung eines in der Praxis bereits vermuteten, aber wenig greifbaren Umstandes. Die Musikpädagogik beschäftigt sich ja auch seit Längerem mit dem Umgang mit Heterogenität (vgl. Vogt, 2012b). Wir können jetzt aber genau abschätzen, wie stark die Leistungsheterogenität im Musikunterricht tatsächlich ist, und dies in Zahlen ausdrücken und kommunizieren.

Die zweite Frage dieses Beitrags beschäftigt sich mit den mutmaßlich verantwortlichen Hintergrundvariablen für diese Leistungsheterogenität. Es zeigte sich, dass bezüglich des Geschlechts Mädchen in Klasse 9 generell höhere musikpraktische Leistungen zeigen als Jungen. Dies kann beispielsweise darauf zurückgeführt werden, dass das Fach Musik gerade in der Altersgruppe der 15–16-Jährigen eher als ein „Mädchenfach“ angesehen wird und dass die Jungen dieses Fach mehrheitlich im Zuge ihrer Identitätsfindung vernachlässigen (vgl. Hess, 2015). Selbstverständlich gibt es männliche Schüler, die diesem Muster nicht entsprechen; sie finden aber in den referierten Standardabweichungen durchaus ihren Platz.

Bei der Analyse der Unterschiede zwischen den Schulformen ergab sich wenig überraschend, dass Schüler musischer Gymnasien in allen drei Dimensionen musikpraktischer Kompetenzen erheblich überlegen waren. Da vermutlich Eltern instrumental begabter bzw. geförderter Kinder diese eher an einem musischen Gymnasium anmelden als an einem normalen Gymnasium, war zu erwarten, dass gerade in dieser Schulform besonders viele Schüler überdurchschnittliche Testergebnisse erzielen würden. Darüber hinaus haben Schüler an musischen Gymnasien erheblich mehr Musikunterricht als gleichaltrige Schüler an anderen Schulformen. Die höheren Leistungen könnten also auch durch die erhöhte fachliche Beschulung begründet werden. Für diesen Grund spricht, dass sich kaum Unterschiede zwischen Gymnasiasten und Realschülern finden lassen. Lediglich im Singen schneiden die Gymnasiasten besser ab, nicht jedoch im Instrumentalspiel und beim rhythmischen Musizieren. Folgt man der Argumentation, dass mehr Übung bessere Leistungen (und damit höhere Kompetenzniveaus) erzeugt, so müsste man annehmen, dass an Gymnasien im Musikunterricht mehr gesungen wird als an Realschulen, dann ließen sich die Ergebnisse mit diesem einfachen Grundsatz erklären. Ob es sich tatsächlich so verhält, kann mit den hier vorliegenden Daten allerdings nicht beantwortet werden.

Interessant sind auch die Analysen zum Einfluss des außerschulischen Instrumentalunterrichts. Die generelle Überlegenheit der Instrumentalschüler um mehr als eine Standardabweichung lässt zunächst vermuten, dass dieser Zusatzunterricht tatsächlich die Hauptursache für die gefundene Leistungsheterogenität darstellt. Die nach Unterrichtsjahren unterteilte Detailanalyse zeigt, dass tatsächlich diejenigen, die bereits länger Instrumentalunterricht erhielten, in allen drei Dimensionen höhere Leistungen erzielen. Die potentiell mögliche Erklärung, dass gerade die musikpraktisch höher begabten Kinder auch in früherem Alter mit dem Instrumentalunterricht beginnen und weniger hoch begabte Kinder entsprechend später, scheint aber nur auf den ersten Blick plausibel. Wäre dies der Fall, müssten tatsächlich die gefundenen Standardabweichungen der einzelnen Subgruppen erheblich kleiner als beobachtet ausfallen, weil zumindest die Gruppen der Instrumentalschüler deutlich homogener sein sollten. Dass aber (mit nur wenigen Ausreißern) die Standardabweichungen in allen fünf Dauerngruppen annähernd gleich sind, spricht eher dafür, dass der außerschulische Instrumentalunterricht tatsächlich Quelle musikpraktischer Leistungsunterschiede und nicht Ergebnis begabungsorientierter Selektionsprozesse ist.

Über das tatsächliche Ausmaß des Einflusses dieses zusätzlichen Unterrichts können jedoch auf der hier vorgestellten Datengrundlage keine gesicherten Aussagen getroffen werden. Das gewählte Raster in Zwei-Jahres-Schritten ist eindeutig zu grob, um hier mit statistischen Methoden auf belastbare Annäherungen zu kommen. Die Ergebnisse der Regressionsanalysen können daher eher als eine untere Abschätzung des wahren Einflusses angesehen werden und scheinen aus diesem Grund auf einen tatsächlich bedeutsamen Einfluss des außerschulischen Instrumentalunterrichts hinzudeuten. Aber selbst wenn genug Daten vorlägen, um ein erheblich feineres Raster anzulegen, bliebe fraglich, ob die Dauer des Instrumentalunterrichts tatsächlich als Operationalisierung ausreichend ist. Vielmehr müsste die Qualität des Instrumentalunterrichts, die Intensität der Beschäftigung mit dem Instrument oder die Effektivität des Übens hier berücksichtigt werden. Bislang können wir all diese Konstrukte allerdings noch nicht valide unter forschungsökonomisch vertretbarem Aufwand in größeren Gruppen messen.

Insgesamt kann festgehalten werden, dass die Teilnahme am außerschulischen Instrumentalunterricht zumindest zum Teil dafür verantwortlich ist, dass im Musikunterricht eine große Leistungsheterogenität in den Dimensionen musikpraktischer Kompetenzen anzufinden ist. Mit dieser Tatsache arbeiten Musiklehrer seit Jahrzehnten und haben vermutlich adaptive Strategien entwickelt, damit zurechtzukommen, oder diese Leistungsdifferenzen sogar produktiv zu nutzen (z. B. Terhag & Winter, 2011). Um ein genaueres Bild der Ursachen zu bekommen, ist allerdings weitere Forschung wünschenswert. Und ob diese Heterogenität der Kompetenzen in anderen Fächern mit außerschulischer Förderung (z. B. Sport, Kunst) ebenso groß ist bzw. die Leistungsbreiten in Fächern wie Physik, Geschichte oder Deutsch geringer, wird weitere Forschung aufdecken müssen.

Literatur

- Autorengruppe Bildungsberichterstattung (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann.
- Budde, J. (2012). Problematisierende Perspektiven auf Heterogenität als ambivalentes Thema der Schul- und Unterrichtsforschung. *Zeitschrift für Pädagogik*, 4, 522–540.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson.
- Gröhlich, C., Scharenberg, K. & Bos, W. (2009). Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus? *Journal for educational research online*, 1(1), 86–105.
- Hasselhorn, J. (2015). *Messbarkeit musikpraktischer Kompetenzen von Schülerinnen und Schülern – Entwicklung und empirische Validierung eines Kompetenzmodells*. Münster: Waxmann.
- Hasselhorn, J. & Grollmisch, S. (2014). Die Colored Music Grid (CMG) App – Ein neues Eingabeinterface zur Erfassung instrumenten-unabhängiger instrumentaler Leistungsfähigkeit. In W. Auhagen, C. Bullerjahn & R. v. Georgi (Hrsg.), *Offenohrigkeit. Ein Postulat im Fokus* (S. 231–234). Göttingen: Hogrefe.
- Hasselhorn, J. & Lehmann, A. C. (2014). Entwicklung eines empirisch überprüfbareren Modells musikpraktischer Kompetenz (KOPRA-M). In B. Clausen (Hrsg.), *Teilhabe und Gerechtigkeit* (S. 77–94). Münster: Waxmann.
- Hess, F. (2015). Das ist doch nichts für echte Kerle! Zum Zusammenhang zwischen Geschlechtsrollen-Selbstbild und Einstellungen zum Musikunterricht. Ergebnisse der Studie „Musikunterricht aus Schülersicht“. In A. Bartsch & J. Wedl (Hrsg.), *Teaching Gender? Geschlecht in der Schule und im Fachunterricht*. Bielefeld: transcript.
- Hoene, S. & Thurmann, B. (2011). *Umgang mit Heterogenität im Musikunterricht*. Kronshagen: IQHS.
- Jordan, A. K. (2014). *Empirische Validierung eines Kompetenzmodells für das Fach Musik: Teilkompetenz „Musik wahrnehmen und kontextualisieren“*. Münster: Waxmann.
- Knigge, J. (2014). Der Kompetenzbegriff in der Musikpädagogik: Verwendung, Kritik, Perspektiven. In J. Vogt, M. Brenk & F. Heß (Hrsg.), *(Grund-)Begriffe musikpädagogischen Nachdenkens – Entstehung, Bedeutung, Gebrauch* (S. 105–135). Münster: LIT.
- Moosbrugger, H. (2008). Item-Response-Theorie (IRT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 215–259). Heidelberg: Springer.
- Niessen, A., Lehmann-Wermser, A., Knigge, J. & Lehmann, A. C. (2008). Entwurf eines Kompetenzmodells ‚Musik wahrnehmen und kontextualisieren‘. *Zeitschrift für Kritische Musikpädagogik, Sonderedition: Bildungsstandards und Kompetenzmodelle für das Fach Musik?*, 3–33.
- Persky, H. R., Sandene, B. A. & Askew, J. M. (1997). *The NAEP 1997 Arts Report Card: Eighth-Grade Findings from the National Assessment of Educational Progress*. Washington: ED Pubs.
- Reh, S. (2005). Warum fällt es Lehrerinnen und Lehrern so schwer, mit Heterogenität umzugehen? Historische und empirische Deutungen. *Die Deutsche Schule*, 1, 76–86.
- Reinders, H., Ditton, H., Gräsel, C. & Gniewosz, B. (Hrsg.) (2011). *Empirische Bildungsforschung*. Wiesbaden: VS.

- Schäfer, T. (2011). *Statistik II. Inferenzstatistik*. Wiesbaden: VS.
- Terhag, J. & Winter, J. K. (2011). *Live-Arrangement: vom Pattern zur Performance*. Mainz: Schott.
- Tillmann, K. J. (2007). Viel Selektion – wenig Leistung. Ein empirischer Blick auf Erfolg und Scheitern in deutschen Schulen. In D. Fischer & V. Elsenbast (Hrsg.), *Zur Gerechtigkeit im Bildungssystem* (S. 25–37). Münster: Waxmann.
- Vogt, J. (2012a). Einleitung: Vom Umgang der Musikpädagogik mit Heterogenität. In J. Vogt (Hrsg.), *Musikpädagogik und Heterogenität. Sitzungsbericht 2011 der Wissenschaftlichen Sozietät Musikpädagogik* (S. 6–19). Münster: LIT.
- Vogt, J. (2012b). *Musikpädagogik und Heterogenität. Sitzungsbericht 2011 der Wissenschaftlichen Sozietät Musikpädagogik*. Münster: LIT.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.

Johannes Hasselhorn
Technische Universität Dortmund
Institut für Musik und Musikwissenschaft
Emil-Figge-Straße 50
44227 Dortmund
johannes.hasselhorn@tu-dortmund.de

Andreas C. Lehmann
Hochschule für Musik
Hofstallstr. 6–8
97084 Würzburg
ac.lehmann@hfm-wuerzburg.de