

# Determinanten und Verläufe musikalischer Begabung im Jugendalter

Philippe Labonde  · Daniel Müllensiefen

Eingegangen: 9. August 2021 / Überarbeitet: 31. März 2022 / Angenommen: 30. Mai 2022  
© Der/die Autor(en) 2022

**Zusammenfassung** Nur wenige Langzeitstudien beschreiben bislang musikalische Entwicklungsverläufe mit einem Fokus auf Begabungsforschung (Müllensiefen 2017). Diese Lücke will der vorliegende Beitrag schließen und anhand der Daten aus einer internationalen Langzeitstudie die Frage beantworten, wie hilfreich unterschiedliche Begabungsansätze für die Modellierung musikalischer Entwicklung sind. Dabei soll untersucht werden, wie sich verschiedene Definitionen musikalischer Begabung und deren statistische Umsetzung auf die Identifikation musikalisch Begabter auswirken. Hierfür werden klassische Begabungskriterien nach Gagné (1998) und der TAD-Ansatz (Preckel et al. 2020) zur Klassifizierung der Studienteilnehmer/innen verwendet und die Ergebnisse miteinander verglichen. In beiden Ansätzen erweisen sich fluide Intelligenz, Arbeitsgedächtniskapazität, die Persönlichkeitsdimension Offenheit und der Glaube an die Veränderbarkeit von Musikalität durch Training als robuste Prädiktoren. Als zentraler Unterschied zeigt sich, dass die perzeptuellen musikalischen Fähigkeiten der Begabten, die nach traditionellen Kriterien ausgewählt wurden, ab dem ersten Messzeitpunkt sinkt. Die Ergebnisse werden mit Referenz auf das Phänomen der Regression zur Mitte diskutiert und legen nahe, dass die Analyse von Langzeitdaten unter Berücksichtigung des TAD-Ansatzes einen geeigneteren Ansatz für Entscheidungen über die musikalische Begabung von Jugendlichen darstellen.

**Schlüsselwörter** Wachstumsmodelle · Langzeitstudie · Musik · Begabungsentwicklung

---

Philippe Labonde · Prof. Dr. Daniel Müllensiefen  
Hanover Music Lab, Hochschule für Musik, Theater und Medien Hannover, Neues Neues  
Haus 1, 30175 Hannover, Deutschland  
E-Mail: philippelabonde@gmx.de

Prof. Dr. Daniel Müllensiefen (✉)  
Goldsmiths, University of London, New Cross Road, New Cross, London, SE14 6NW, Großbritannien  
E-Mail: d.mullensiefen@gold.ac.uk

## Determinants and trajectories of musical talent in adolescence

**Abstract** So far, only a few long-term studies have described musical developmental trajectories with a focus on giftedness research (Müllensiefen 2017). This article aims to fill this gap and to answer the question of how helpful different approaches to giftedness are for modeling musical development using data from an international long-term study. The aim is to investigate how different definitions of musical giftedness and their statistical implementation affect the selection of musically gifted individuals. For this purpose, classical giftedness criteria according to Gagné (1998) and the TAD approach (Preckel et al. 2020) are used for classification and the results are compared. In both approaches, fluid intelligence, working memory capacity, the personality dimension openness, and the belief in the changeability of musicality through training, prove to be robust predictors. As a central point of difference, the assessed perceptual musical abilities of gifted participants, selected according to classical criteria, decrease from the first measurement time point. The results are discussed with reference to the phenomenon of regression to the mean, suggesting that long-term data and the TAD approach provide a more suitable approach for making decisions about musical giftedness in adolescents.

**Keywords** Growth modeling · Longitudinal study · Music · Talent development

### 1 Einleitung

Die Identifizierung von überdurchschnittlicher musikalischer Leistung und Begabung bei Kindern und Jugendlichen hat eine lange Tradition in der Musikpädagogik und Musikpsychologie (vgl. Seashore 1919). Denn musikalische Begabung frühzeitig zu identifizieren, ist eine der wichtigsten Voraussetzungen für die Förderung musikalischen Potentials und das Erreichen herausragender musikalischer Leistungen. Dennoch existieren derzeit keine etablierten Verfahren zur Messung musikalischer Begabung, so dass bei Förderentscheidungen und Auswahlprozessen bspw. die Zusammensetzung oder auch die (nicht-) Orientierung an standardisierten Kriterien von lokalen Jurys eine große Rolle spielen können. Besonders problematisch ist das Fehlen verlässlicher Diagnoseverfahren für Kinder und Jugendliche aus bildungsfernen und sozial benachteiligten Bereichen der Gesellschaft. Denn wenn überhaupt, dann findet die Begegnung mit Musik hier oft abseits der traditionellen musikalischen Bildungseinrichtungen (Privatunterricht, Schulorchester, Musikschule) statt, die üblicherweise die Identifizierung und Förderung von musikalischer Begabung übernehmen. Das Spielen eines Instrumentes ist so in der Regel Voraussetzung für das Entdecken von Begabung, was auch bedeutet, dass Kinder mit hohem musikalischem Potenzial, aber ohne Instrumentalunterricht, übersehen werden können. Zudem liegt der Fokus bei traditionellen Ansätzen zur Identifizierung musikalischer Begabung häufig auf den Leistungen, die in einem bestimmten Alter erbracht werden. Dass diese Leistungen jedoch als Ergebnis eines Entwicklungsprozesses zu verstehen sind, machen neuere Ansätze der allgemeinen Begabungsforschung deutlich (vgl. Subotnik et al. 2011, 2012). Auch in der musikpädagogischen Forschung

fehlt oft der Fokus auf den musikalischen Entwicklungsverlauf, so dass psychologische Faktoren, die musikalische Entwicklung entscheidend beeinflussen können, in der Regel nicht bei der Diagnose von Begabung berücksichtigt werden.

Deshalb ist es das Ziel dieses Beitrags auf Grundlage eines aktuellen Begabungsmodells – dem Talent Development in Achievement Domains (TAD, vgl. Preckel et al. 2020) – aufzuzeigen, wie musikalische Begabung ohne Berücksichtigung der Fähigkeiten am Instrument und mit einem Fokus auf die Entwicklung von Fähigkeiten über die Zeit mit Hilfe quantitativer Modelle beschrieben werden kann. Dieser am TAD-Modell angelehnte neue Ansatz soll zudem mit traditionellen Ansätzen verglichen werden, die musikalische Begabung mit hoher Leistung oder hohen Fähigkeiten zu einem bestimmten Zeitpunkt gleichsetzen. Zuletzt werden kognitive und psychosoziale Faktoren als Proxyvariablen in die Analyse miteinbezogen.

## 2 Theorieansätze zu musikalischem Potenzial und musikalischen Fähigkeiten

Ein in der Musikpädagogik sehr einflussreiches Begabungskonzept ist Gagnés „Integrative Model of Talent Development“ (IMTD), welches von McPherson und Williamon (2016) für den musikalischen Bereich spezifiziert wurde. In diesem Modell wird zwischen „Gifts“ und „Talents“ unterschieden. Ersteres bezeichnet „natürliche“ Fähigkeiten, die nicht durch systematisches Training erworben werden und die in der Literatur meist als Potenzial oder Begabung benannt werden. Dabei gelten laut Gagné (Gagné und McPherson 2016, S. 5) die besten 10% einer Altersgruppe zu der Kohorte der begabten Individuen. Als „Talents“ werden in diesem Modell systematisch entwickelte Fähigkeiten oder Fertigkeiten bezeichnet<sup>1</sup>. Talentierte Individuen sind demnach diejenigen, die (bei gleicher Lernzeit) die obersten 10% an Leistungen abbilden. Als musikspezifische Fähigkeiten benennen McPherson und Williamon (2016) das Spielen, Improvisieren, Komponieren, Arrangieren, Analysieren, Dirigieren und die Wertschätzung von Musik. Für den Entwicklungsprozess, in dem das Potenzial in systematisch entwickelte Fähigkeiten umgesetzt wird, benennt Gagné eine Reihe von Katalysatoren. Katalysatoren sind Eigenschaften oder Umstände, die dabei helfen den Umwandlungsprozess von „gifts“ zu „talents“ zu beschleunigen oder im negativen Fall Begabte davon abhalten ihre Leistungsfähigkeit voll auszuschöpfen. Die Konstruktion eines quantitativen Modells, durch Welches

---

<sup>1</sup> Im Deutschen wird in der differentiellen Psychologie wie auch in der Musikpsychologie oft zwischen den Begriffen Fähigkeiten und Fertigkeiten unterschieden. Dabei bezeichnen Fähigkeiten in der Regel psychologische Konstrukte, die als nur gering durch spezifische Umwelteinflüsse veränderbar gelten (z. B. Intelligenz). Im Gegensatz wird der Begriff Fertigkeiten für Konstrukte verwendet, von denen angenommen wird, dass sie zu einem erheblichen Anteil lern- und förderbar sind. Zur Differenzierung der Terminologie in der Musikpsychologie s. a. Müllensiefen und Hemming (2018). Dennoch wird im vorliegenden Beitrag durchgehend der gebräuchlichere Begriff Fähigkeiten verwendet, weil die vorgestellten Theorieansätze ihre eigene Terminologie verwenden (gifts, talents bei Gagné und predictors, indicators bei TAD), die nicht uneingeschränkt mit der Unterscheidung in Fähigkeiten v. Fertigkeiten kompatibel ist. Zudem soll eine Apriori-Einschätzung der Veränderbarkeit der gemessenen psychologischen Konstrukte vermieden werden.

der Einfluss der Katalysatoren auf die Entwicklung von musikalischer Begabung modelliert und untersucht wird, steht derzeit jedoch noch aus.

Gagnés Modell wird durch das „Multifactorial Gene–Environment Interaction Model“ (MGIM) von Ullén et al. (2016) ergänzt, welches eine detaillierte Beschreibung verschiedener Genotyp-Umwelt-Interaktionen (passiv, aktiv, evokativ) darstellt, die auf Plomin et al. (1977) zurückgehen. Durch dieses Modell wird Gagnés Annahme, dass genetische Komponenten bei der Talententwicklung von großer Bedeutung sind, erweitert und in Hinsicht auf genetische Prozesse spezifiziert. Diese differenzierte Ausarbeitung verschiedener Arten von Genotyp-Umwelt-Interaktionen scheint auch für den Musikbereich sinnvoll und gut anwendbar. Eine passive Genotyp-Umwelt-Interaktion läge beispielsweise dann vor, wenn Kinder von ihren Eltern Fähigkeiten vererbt bekommen, die für das Musikmachen bedeutsam sind und sie gleichzeitig durch den Einfluss ihrer musikalischen Eltern in einer Umgebung aufwachsen, die reich an musikalischen Möglichkeiten ist. Dagegen läge eine aktive Interaktion dann vor, wenn der Genotyp beeinflusst, ob und welcher musikalischen Beschäftigung nachgegangen wird. Das kann zum Beispiel das Auslösen des eigenen neuronalen Belohnungssystems sein, wenn durch die motorische und kognitive Veranlagung Instrumente leicht und schnell erlernt werden. Hierdurch erfahrene externe oder eigene Bestätigung kann wiederum zu dem Ansporn führen, weiterhin Musik zu machen bzw. sogar noch mehr persönliche Ressourcen in Musik zu investieren. Schließlich sind evokative Genotyp-Umwelt-Einflüsse dadurch gekennzeichnet, dass Individuen mit ausgeprägten musikalischen Anlagen Musiklehrer/innen oder andere musikalische Peers auf sich aufmerksam machen, und so ggf. die Qualität und Quantität der musikalischen Beschäftigung steigt.

Durch die Ausdifferenzierung der möglichen Interaktionen zwischen genetischen Anlagen und der Umwelt gehen Ullén et al. (2016) auch auf die Bedeutung des gezielten Übens („deliberate practice“) ein, welche zudem in der Expertiseforschung eine wichtige Rolle spielt. Obwohl das IMTD und das MGIM exzellente Ansätze darstellen, die die Komplexität von musikalischer Talententwicklung verständlich machen, bleibt es weiterhin schwierig ihre modellhaften Annahmen und Mechanismen in quantitative und empirisch überprüfbare Modelle der Begabungsdiagnostik und -entwicklung umzusetzen.

Gemeinsam ist diesen Ansätzen, dass herausragende musikalische Fähigkeiten als hinreichender Indikator für Begabung verwendet werden. Gagné (1998) schlägt dabei vor, Subkategorien von begabten und talentierten Individuen anhand von unterschiedlichen Schwellenwerten zu unterscheiden (s. a. Gagné und McPherson 2016). Die Individuen, die bei der Verteilung musikalischer Fähigkeiten in die obersten 10 % fallen klassifiziert Gagné als „milde“ Hochbegabung (vgl. Gagné 1998). Dabei nutzt Gagné bewusst Prozentwerte und kein an die Standardabweichung angelehntes Kriterium, damit diese Schwellenwerte auch bei nicht normierten Tests oder nicht normalverteilten Merkmalen angewendet werden können. In ihrem kurzen Review zu Schwellenwerten bei Intelligenztests zur Erfassung von kognitiv hochbegabten Individuen bemerken McIntosh et al. (2018, S. 601), dass die meisten Studien einen Schwellenwert von mehr als zwei Standardabweichungen über dem Mittelwert der Normstichprobe annehmen. Dabei raten McIntosh et al. aber dringend davon ab, den Ein- oder Ausschluss von leistungsstarken Individuen in Fördermaßnahmen allein

von einem Wert abhängig zu machen und raten dazu, auch weitere Faktoren sowie die individuelle Entwicklung in Betracht zu ziehen.

Ein neuer und in dieser Hinsicht vielversprechender Ansatz, mit dem Potenziale und Fähigkeiten musikalisch leistungsstarker Kinder und Jugendlicher abgebildet werden können, ist der „Talent Development in Achievement Domains“ (TAD) genannte Ansatz von Preckel et al. (2020), der auch für den Bereich der Musik ausgearbeitet wurde. Gemäß des TAD-Ansatzes ist musikalische Begabung oder musikalisches Potenzial eine latente Variable, die nicht direkt beobachtet werden kann. Allerdings wird postuliert, dass sich die latente Variable Begabung durch Proxyvariablen abschätzen lässt. Proxyvariablen werden als psychologische Faktoren definiert (z. B. Persönlichkeit, kognitive und Wahrnehmungsfähigkeiten, Selbstkonzepte, motivationale und psycho-soziale Variablen), die in der Interaktion mit geeignetem musikalischem Input (Trainingsintensität und -qualität z. B. beim instrumentalen Üben, Singen, aktiven Hören) die Vorhersage zukünftiger musikalischer Leistung erlauben. Der TAD-Ansatz schlägt auf Grundlage der existierenden Literatur konkrete Proxyvariablen vor. Dabei wird angenommen, dass für verschiedene Entwicklungsstadien, Altersstufen, Instrumente und musikalische Genres unterschiedliche Sets von Proxyvariablen bedeutsam sein können. Das heißt die Auswahl und Gewichtung der Faktoren, die den Entwicklungsverlauf beeinflussen sind nicht statisch, sondern verändern ihren Einfluss im Laufe der Entwicklung. Beispielsweise benötigt ein/e professionelle/r Musiker/in für eine erfolgreiche Karriere auch kommunikative und soziale Fertigkeiten, um mit Kolleg/innen, Publikum oder Agent/innen erfolgreich zu kommunizieren. Dagegen muss sich ein/e junge/r Musiker/in deutlich intensiver auf die Entwicklung motorischer Fertigkeiten konzentrieren. Um dieser Verschiebung von Einflussgrößen auf die musikalische Entwicklung Rechnung zu tragen, werden im TAD-Modell vier Level der Talententwicklung spezifiziert, die Müllensiefen et al. (2022) für den Bereich der Musik auch inhaltlich bestimmt haben:

1. **Aptitude:** In dieser Phase der Entwicklung werden vorhandene musikalische Veranlagungen in rudimentäre musikalische Fähigkeiten umgesetzt, indem die ersten Stationen systematischer Beschäftigung mit Musik in der Schule oder beim selbstregulierten Lernen durchlaufen werden. Hier zählen vor allem allgemeine kognitive Fähigkeiten (z. B. Gedächtnisleistung, Intelligenz), bestimmte Persönlichkeitsdimensionen (z. B. Offenheit für Neues) und motivationale Faktoren, intuitive Hörwahrnehmungs- und Produktionsleistungen sowie emotionales Verständnis von Musik zu den bedeutsamen Prädiktoren musikalischer Leistung.
2. **Competence:** In der Competence-Phase werden verschiedene musikbezogene Fähigkeiten erlernt, die im Allgemeinen die Bandbreite an Repertoire, Ausdruck und Kontexten erweitern, in denen sich die Individuen ausdrücken können. Dazu gehören je nach Genre das Verständnis zunehmend komplexerer musikalischer Strukturen, der Ausdruck von musikalischen Emotionen und die Regeln für das Zusammenspiel mit Anderen. In der Competence-Phase sind es vor allem die Fähigkeit selbstständig und zielgerichtet zu lernen, musikbezogene Einstellungen (z. B. Selbstkonzept, „growth mindset“) und musikalische Leistungsbereitschaft, die als Prädiktoren für die weitere Entwicklung von musikalischen Fähigkeiten angenommen werden.

3. Expertise: Hier werden musikalische Fähigkeiten erreicht, die von der Peer-Group anerkannt werden und eine berufliche Karriere im Musikbereich ermöglichen. Als Prädiktoren des Entwicklungsverlaufes in dieser Phase kommen allgemeine psychologische Faktoren wie zum Beispiel Resilienz, Offenheit und die Fähigkeit zur konstruktiven Kritik des eigenen Schaffens in Frage.
4. Transformational achievement: Diese höchste Entwicklungsstufe ist durch Erfolge, Anerkennung und Einfluss in der Musikwelt gekennzeichnet. Konzepte wie persönliches Charisma oder kreative Produktivität stehen hier als Marker für die erfolgreiche Entwicklung in dieser Phase.

Eine wichtige Implikation des TAD-Ansatzes ist die Minimierung des „selection bias“ (vgl. Geddes 1990), der bei der Untersuchung musikalisch Hochbegabter meist dadurch entsteht, dass ausschließlich musikalisch leistungsstarke Kinder und Jugendliche untersucht werden, die schon eine intensive Förderung im Bereich der westlichen Kunstmusik genossen haben (vgl. Gembris 2002). Die Gefahr des „selection bias“ besteht darin, dass post-hoc Variablen als Prädiktoren musikalischer Leistungsfähigkeit definiert werden, ohne dabei zu berücksichtigen, welchen Einfluss diese Variablen auf die Aufnahme und Fortführung musikalischen Trainings hatten und somit welche Bedeutung diesen Variablen für die Feststellung musikalischen Potentials a-priori zukommt.

Ein zusätzlicher wichtiger Aspekt des TAD-Ansatzes ist die Entkoppelung von Alter und Entwicklungsstufe, was die Möglichkeit von unterschiedlichen Entwicklungsverläufen in verschiedenen musikalischen Genres eröffnet. So ist ein später Einstieg in das Instrumentalspiel in der populären Musik keine Seltenheit, wohingegen in der klassischen Musik systematisches Training im frühen Kindesalter in der Regel eine Voraussetzung für eine spätere Karriere ist (vgl. Creech et al. 2008). Empirische Modelle könnten eine entsprechende Differenzierung von Alter, musikalischem Genre und Entwicklungsstufe berücksichtigen.

Für den vorliegenden Beitrag ist es für die empirische Identifizierung von Begabung gemäß des TAD-Ansatzes wichtig, dass Lernfortschritt innerhalb eines Zeitintervalls unter ansonsten vergleichbaren Bedingungen im Hinblick auf andere Personen gemessen wird. Das heißt, begabte Schülerinnen und Schüler (SuS) zeigen bei identischem musikalischen Input im selben Zeitraum einen größeren Leistungszuwachs als unbegabte SuS. Musikalischer Input könnte beispielsweise über die Anzahl der Unterrichtsstunden bei derselben Lehrperson und die Anzahl der Übestunden definiert werden. Unterschiede in der individuellen Begabung werden als Ursache für die Unterschiede in den Leistungszuwächsen angenommen. Begabung ist zwar nicht direkt beobachtbar und wahrscheinlich zumindest teilweise genetisch determiniert. Aber nach TAD ist es möglich, eine Auswahl von Proxyvariablen zu identifizieren, die die Funktion der latenten Variable Begabung näherungsweise beschreibt.

Das Kriterium des Leistungszuwachses über die Zeit, dass für den TAD-Ansatz zentral ist, steht im Gegensatz zu absoluten Kriterien der musikalischen Leistung (z. B. top 10 % oder top 1 % der musikalischen Leistungen in einer Altersgruppe), die in anderen musikalischen Begabungsmodellen vorherrschen (z. B. Gagné und McPherson 2016). Zwar beschreiben auch Gagné und McPherson (2016, S. 23) ein Zuwachskriterium („musical progress quotient“), bleiben aber eine empirische

Implementierung schuldig. Zudem fehlt dort die Differenzierung zwischen der Zuwachsrate als messbare Größe und der Begabung als latentes Konstrukt, welches durch Proxyvariablen näherungsweise beschrieben und für die empirische Modellierung verwendet werden kann.

## 2.1 Messung musikalischer Fähigkeiten

Die Fähigkeit musikalische Strukturen zu differenzieren und subtile akustisch-musikalische Unterschiede wahrzunehmen wurde schon von Michaelis (1805) und Galton (1869) als fundamental für die menschliche Musikalität angesehen. Galtons Überlegungen beeinflussten dann die Entwicklung der ersten musikalischen Testbatterie von Seashore (1919), bei der die jugendlichen Testpersonen (10–19 Jahre) Unterschiede zwischen kurzen Stimuli auf sechs grundlegenden auditorischen Dimensionen (Tonhöhe, Lautstärke, Rhythmen etc.) erkennen müssen. Auch die musikalischen Testbatterien von Wing (1939), Bentley (1966) und Gordon (1989, 1995) und vielen anderen zielen auf die Unterscheidungsfähigkeit für grundlegende musikalisch-akustische Parameter ab. Verwendet werden dafür fast ausschließlich abstrakte und relativ artifizielle Stimuli. Diesem Ansatz der Testkonstruktion liegt die Annahme zugrunde, dass die getesteten basalen Wahrnehmungsfähigkeiten nur wenig von musikalischem Training und Erfahrung beeinflusst werden und es so möglich sei, musikalische Begabung unverfälscht zu messen (Seashore verwendet den Begriff „talents“, Gordon „aptitude“).

Auch die Mehrzahl der in den letzten Jahren veröffentlichten modernen musikalischen Testbatterien (z. B. Law und Zentner 2012; Ullén et al. 2014; Wallentin et al. 2010) halten an dem traditionellen Ansatz fest, einfache und relativ abstrakte, d. h. von konkreter Musik losgelöste Fähigkeiten zu testen. So bleiben viele Facetten menschlicher Musikalität, die sich beim Umgang mit echter Musik zeigen, unberücksichtigt (vgl. die Kritik von Gembris 2002; Karma 1984; Sloboda und Howe 1991). Zudem sind auch viele moderne Testbatterien weiterhin stilistisch an die europäische Kunstmusik und die klassische Gehörbildung angelehnt und können so jenen Kindern einen Vorteil verschaffen, die Training in diesem Bereich erhalten haben. Schließlich steht eine umfassende Überprüfung des prognostischen Wertes der modernen Testbatterien für die Identifizierung von musikalischer Begabung derzeit noch aus. Viele der älteren traditionellen Testbatterien sind dafür kritisiert worden, dass sie generell nur einen geringen prognostischen Wert besitzen (vgl. Hodges und Haak 1996; Manturzevska 1990; Norton et al. 2005; Winner und Martino 2000).

Einen komplementären Ansatz verfolgt der Goldsmiths Musical Sophistication Index (Gold-MSI, vgl. Müllensiefen et al. 2014). Der Gold-MSI Selbstauskunftsfragebogen erhebt neben der Ausprägung musikalischer Wahrnehmungsfähigkeiten auch den produktiven (Singen, aktiver Umgang mit Musik, musikalisches Training) und emotionalen Umgang mit Musik. Besonderer Wert wird bei dem Gold-MSI u. a. darauf gelegt musikalische Umgangsweisen zu berücksichtigen, die für musikalische Stile auch abseits der westlichen Kunstmusik bedeutsam sind. Ergänzend zum Gold-MSI Selbstauskunftsfragebogen ist in den letzten Jahren eine Batterie neuer Hörtests entwickelt und veröffentlicht worden, die moderne psychometrische Verfahren verwenden (Item Response Theory, Computerised Adaptive Testing, Automatic Item

Generation), um sich dem Fähigkeitsniveau der Versuchspersonen adaptiv anzupassen und so eine maximale Testeffizienz zu erreichen. Der adaptive Mechanismus der Tests und die Verwendung einer einheitlichen Fähigkeitskala macht es möglich die Tests über längere Entwicklungszeiträume und für Stichproben mit sehr heterogenen Fähigkeitsniveaus zu verwenden. Gleichzeitig verwenden die im Zuge des Gold-MSI entwickelten Tests so weit wie möglich realistische und komplexe musikalische Stimuli aus vielen unterschiedlichen Genres, die der musikalischen Lebenswirklichkeit einer breiten Öffentlichkeit in westlichen Gesellschaften entsprechen. Zu den Tests, die mit Hilfe dieses Paradigmas entwickelt wurden, gehören der Melodic Discrimination Test (MDT; vgl. Harrison et al. 2017), der Mistuning Perception Test (MPT; vgl. Larrouy-Maestri et al. 2019), der Computerised Adaptive Beat Alignment Task (CA-BAT; vgl. Harrison und Müllensiefen 2018) und der Musical Emotion Discrimination Task (EDT; vgl. MacGregor und Müllensiefen 2019)<sup>2</sup>. Die Verwendung der Gold-MSI-Tests in dieser Studie soll zum einen dem selection-bias entgegenwirken, da ausschließlich Hörfähigkeiten erhoben werden, die auch ohne vorherigen Instrumentalunterricht oder andere formale musikalische Vorkenntnisse entwickelt werden können. Zum anderen vermeiden die Tests durch die Verwendung realistischer musikalischer Stimuli und Höraufgaben den einseitigen Fokus auf einen bestimmten musikalischen Stil oder eine bestimmte Art musikalischer Ausbildung.

## 2.2 Auswahlkriterien und Förderpraktiken musikalisch leistungsfähiger Jugendlicher

Obwohl in Deutschland allgemeinbildende Schulen mit ausgeprägtem Musikprofil existieren, findet die Förderung musikalisch leistungsstarker Individuen in Deutschland i. d. R. außerhalb des schulischen Unterrichts statt (vgl. Clausen und Lessing 2018, S. 392). Hierfür stehen studienvorbereitende Kurse an Musikschulen und Frühförderungsprogramme an Musikhochschulen zur Verfügung, wie beispielsweise die Pre-Colleges der Musikhochschulen in Köln und Würzburg sowie das Institut für Frühförderung in Hannover. Für die Aufnahme in Frühförderungsprogramme und das Jungstudium an Musikhochschulen müssen Auswahlverfahren absolviert werden, bei denen vor allem herausragende individuelle Fähigkeiten im Instrumentalspiel und oft auch Prüfungen in Musiktheorie und Gehörbildung gefordert sind (vgl. Hochschule für Musik Würzburg 2017). Eine weitere Möglichkeit zur Auswahl und Förderung musikalisch leistungsfähiger Jugendliche sind Jugendensembles, die für viele unterschiedliche Besetzungen und Altersgruppen existieren. Vom Musikinformationszentrum (MIZ) werden derzeit 117 Jugendensembles auf Landes- und 11 auf Bundesebene verzeichnet (vgl. Deutsches Musikinformationszentrum 2021). Aufnahmebedingung sind hinreichende Fähigkeiten im Instrumentalspiel und oft eine Altersbeschränkung. Gefördert werden die Mitglieder dieser Jugendensembles durch intensive Probenarbeit, regelmäßige Auftritte und teilweise auch Kooperationen mit professionellen Ensembles (bspw. bei der Jungen Deutsche Philharmonie

<sup>2</sup> Für eine aktuelle Liste der entwickelten Tests mit Referenzen zu Veröffentlichungen, open source code und Demoanwendungen s. die Demo-Seite des LongGold-Projekts: [https://shiny.gold-msi.org/longgold\\_demo/](https://shiny.gold-msi.org/longgold_demo/).



2021). Eine dritte Säule der Identifikation und Selektion musikalisch leistungsstarker Kinder und Jugendlicher stellen Musikwettbewerbe dar, wie z. B. der bundesweit organisierte Wettbewerb „Jugend musiziert“. Hier erhalten Kinder und Jugendliche die Möglichkeit ihr Instrumentalspiel oder ihre Kompositionen von einer Fachjury einschätzen zu lassen und sich mit Gleichaltrigen zu vergleichen (vgl. Rademacher 2017).

Zusätzlich zu diesen institutionellen Selektions- und Förderinstrumenten, existieren Screening-Fragebögen und Checklisten, die spezifisch für die Identifikation und Selektion von musikalisch leistungsstarken Kindern im Kontext von allgemeinbildenden Schulen konzipiert sind. Über die Verbreitung und praktische Anwendung dieser Selektionswerkzeuge ist jedoch wenig bekannt (vgl. Heller und Perleth 2005; Mayr et al. 2010). Zwar können diese Werkzeuge auch von nicht speziell geschulten Personen eingesetzt werden, jedoch ist die Spezifität dieser Instrumente eher gering (Müllensiefen 2017, S. 138). Sie eignen sich daher eher für das Screening in der Breite und können ggf. mit Hörtests kombiniert werden, die für entsprechende Altersgruppen konzipiert und normiert wurden (s. oben: Bentley 1966; Gordon 1989; Wing 1939).

Insgesamt zeigt sich, dass zumindest in Deutschland in der Praxis keine standardisierten Tests zur Identifikation von musikalischer Begabung zum Einsatz kommen. Eine viel größere Bedeutung kommt dem Instrumentalvorspiel und den Fähigkeiten am Instrument zu, die in einem gewissen Alter erreicht werden. So gut wie keine Bedeutung haben psychologische Faktoren abseits musikalischer Fähigkeiten (z. B. Persönlichkeit, Motivation oder Lerneinstellungen) für die Praxis der Selektion und die gezielte Förderung musikalisch leistungsstarker Kinder und Jugendlicher. Entscheidend ist also meist die subjektive Meinung von Experten/innen und Fachjurs, ob Kinder und Jugendliche in Fördermaßnahmen aufgenommen werden sollen oder nicht. In den wenigen Studien, die sich mit musikalischen Fördermaßnahmen beschäftigen (vgl. Bastian und Koch 2010; Bullerjahn et al. 2020; Manturzewska 1990), wird die Validität, Reliabilität und Objektivität der in der Praxis angewandten Auswahlkriterien meist nicht geprüft. Vielmehr werden die vorgestellten Auswahl- und Fördermaßnahmen der genannten Studien rückwirkend dadurch legitimiert, dass ein Teil der Geförderten später als professionelle Musiker/innen beschäftigt sind. Mit diesem Ansatz ist es aber nicht möglich zu bestimmen, wie viele musikalische begabte Kinder und Jugendliche ‚übersehen‘ werden, d. h. von einem musikalischen Förderprogramm stark profitiert hätten, aber nicht am Auswahlprozess teilnehmen konnten bzw. an den spezifischen Auswahlkriterien gescheitert sind. Ein Problem der meisten Auswahl- und Förderinstrumente ist der vorherrschende Fokus auf westliche Kunstmusik, durch den die Frage nach musikalischer Begabung mit der Frage des musikalischen Geschmacks bzw. der musikalischen Sozialisation eng verknüpft wird und so die Auswahl von musikalischem Talent in anderen musikalischen Bereichen erschwert (vgl. Gembris 2002; Pape 2013). Aus diesen genannten Gründen ist es schwierig, die in der Praxis verbreiteten Auswahlkriterien für musikalisch leistungsstarke Kinder und Jugendliche hinsichtlich ihrer Effektivität und Validität einzuschätzen.

### 2.3 Entwicklung der Forschungsfrage

Die frühe und korrekte Identifizierung von musikalisch leistungsfähigen Individuen ist eine der wichtigsten Voraussetzungen für die zielgerichtete und effiziente Unterstützung zur Entwicklung musikalischen Potentials. In der Forschung wie in der Praxis erweist sich der Mangel an verlässlichen Diagnoseverfahren für Kinder und Jugendliche aus bildungsfernen und sozial benachteiligten Bereichen der Gesellschaft als problematisch. Denn hier findet die Begegnung mit Musik oft abseits der traditionellen musikalischen Bildungseinrichtungen (Musikschule, Privatunterricht, Jugendensemble, Frühförderung an Musikhochschulen) statt, die typischerweise die Identifizierung und Förderung von musikalischer Begabung leisten. Derzeit existiert jedoch kein modernes theoriebasiertes Verfahren zur Identifizierung von überdurchschnittlichem musikalischem Potenzial, welches eine Evidenzbasis für die systematische Entwicklung von herausragenden Leistungen in der Musik liefern könnte. Das Ziel dieses Beitrags ist es deshalb zu überprüfen, ob und inwieweit der vor kurzem entwickelte TAD-Ansatz als Theorierahmen für die Spezifizierung von empirischen Modellen dienen kann, um musikalische Begabungspotentiale zu identifizieren und musikalische Entwicklung vorherzusagen. Dabei soll der TAD-Ansatz mit traditionellen Begabungsansätzen verglichen werden, die sich an statischen Leistungskriterien orientieren. Der Vergleich der Forschungsansätze lässt sich anhand folgender Forschungsfragen gliedern:

Messung perzeptueller musikalischer Leistungsfähigkeit:

1. Wie lässt sich perzeptuelle musikalische Leistungsfähigkeit unabhängig vom Instrumentalspiel und ohne Voraussetzung formaler musikalischer Vorkenntnisse oder Terminologie messen?

Klassische Begabungsansätze:

2. Wie lassen sich, gemäß traditioneller Ansätze, musikalisch Begabte über statische Leistungskriterien identifizieren?
3. Welche extra-musikalischen Variablen sind robuste Prädiktoren für die in 2) identifizierten Begabten?
4. Wie verläuft die Entwicklung musikalischer Fähigkeiten für die in 2) identifizierten Begabten?

TAD-Ansatz:

5. Wie lassen sich Langzeitdaten musikalischer Leistungen zur Identifizierung musikalischer Begabter im Sinne des TAD-Ansatzes verwenden?
6. Welche extra-musikalischen Variablen sind robuste Prädiktoren für die in 5) identifizierten Begabten?
7. Wie verläuft die Entwicklung musikalischer Fähigkeiten für die in 5) identifizierten Begabten?

Vergleich der Ansätze:

8. Wie stark ähneln sich die Ergebnisse des TAD-Ansatzes und der Modelle, die traditionellen Leistungskriterien für die Feststellung individueller Begabung verwenden?

### 3 Methode und Erhebungsinstrumente

Die verwendeten Messinstrumente sind Teil der Langzeitstudie LongGold<sup>3</sup>, welche seit 2015 in England und seit 2018 in Deutschland an Sekundarschulen durchgeführt wird. Ziel des LongGold-Projektes ist es, die wechselseitigen Beziehungen zwischen musikalischer Entwicklung auf der einen Seite und der Entwicklung kognitiver Fähigkeiten, Lerneinstellungen, Persönlichkeit, psychosozialer Faktoren und akademischer Leistungen auf der anderen Seite im Jugendalter zu untersuchen. Für den vorliegenden Beitrag werden aus der LongGold-Datenbank Daten derjenigen Tests und Fragebögen verwendet, die sowohl gemäß traditionellen Ansätzen wie auch nach dem TAD-Ansatz für die Identifizierung und Entwicklung musikalischer Begabung als bedeutsam angenommen werden können (Müllensiefen et al. 2022).

#### 3.1 Erhebung

Die hier verwendeten Daten wurden in einem Zeitraum von 2015 bis 2019 in der LongGold-Studie erhoben. An jeder Schule wurde in diesem Zeitraum einmal im Jahr eine Testung durchgeführt, wobei die Schulen in unterschiedlichen Jahren in die Studie aufgenommen wurden. Die Testungen wurden im Klassenverbund durchgeführt, wobei die Schülerinnen und Schüler (SuS) innerhalb einer Doppelstunde die Online-Testbatterie auf dem eigenen Tablet und mit eigenem Kopfhörer bearbeiteten. Für jede Klasse standen dabei mindestens zwei Testleiter/innen bereit, die die Einhaltung des Testprotokolls gewährleisten und den Teilnehmenden bei Fragen helfen konnten.

#### 3.2 Teilnehmende

Der für die Analysen verwendete Datensatz besteht aus  $N=7451$  Einzeldatensätzen aus dem Erhebungszeitraum, die  $N=4084$  Teilnehmende umfasst. Das durchschnittliche Eintrittsalter im 5. Schuljahr ist  $M=10,5$  mit einer Standardabweichung von  $SD=0,69$ . Von den Teilnehmenden geben 53 % ihr Geschlecht als weiblich, 29,5 % männlich und 1,1 % geben ein non-binäres Geschlecht an. Die restlichen 16,4 % machen keine Angabe zu ihrem Geschlecht. Dabei sind 55,4 % der Stichprobe Teilnehmende aus deutschen und 44,6 % Teilnehmende aus den englischen Sekundarschulen (für weitere Stichprobendaten siehe Tab. 6, 7 und 8 im Anhang).

---

<sup>3</sup> [www.longgold.org](http://www.longgold.org).

### 3.3 Messinstrumente

#### 3.3.1 Abhängige Variablen: Musikalische Fähigkeiten

Im LongGold-Projekt werden musikalische Fähigkeit mit einer Batterie von unterschiedlichen Hörtests erhoben, die kein explizites musikalisches Vorwissen erfordern. Das heißt, es werden weder musikalische Fachterminologie noch instrumentale Fähigkeiten vorausgesetzt, die typischerweise im Musik- oder Instrumentalunterricht vermittelt werden. Die in diesem Rahmen getesteten Hörfähigkeiten können auch durch informelles und inzidentelles Lernen von Musik erworben werden und sind vom präferierten Musikgenre und -stil unabhängig. Dadurch soll insbesondere das Kriterium der Testfairness erfüllt werden, da die musikalischen Hörfähigkeiten auch von Individuen erhoben werden können, die keinen Musik- oder Instrumentalunterricht erhalten und keine oder nur wenig Erfahrung mit westlicher Kunstmusik haben. Alle musikalischen Hörtests verwenden einen adaptiven Mechanismus, der auf einem item response theory-Modell basiert. Dadurch sind alle Hörtests hoch effizient, vermeiden Decken- und Bodeneffekte und ermöglichen eine prinzipielle Abwägung von Testlänge und Messgenauigkeit, welche für den Einsatz in Schulen bedeutsam ist. Eine erste Analyse der Struktur musikalischer Hörwahrnehmungsfähigkeiten zeigt, dass ein Generalfaktor (G-Faktor) aus drei Wahrnehmungstests mit Fokus auf unterschiedlichen musikalischen Dimensionen das latente Konstrukt der musikalischen Hörwahrnehmungsfähigkeiten am besten repräsentiert (Pausch et al. 2022). Dieser G-Faktor wird aus einem Beatwahrnehmungstest (Harrison und Müllensiefen 2018), aus einem Melodieunterscheidungstest (Harrison et al. 2017) und einem Intonationswahrnehmungstest (Larrouy-Maestri et al. 2019) gebildet.

**Computerised Adaptive Beat Alignment Test (CA-BAT)** Der CA-BAT misst die Fähigkeit den zugrunde liegenden Puls eines Musikstückes zu abstrahieren (Harrison und Müllensiefen 2018). Diese Fähigkeit ist fundamental für erfolgreiche musikalische Synchronisation (Honing 2012). Konzeptionell verwandt ist der CA-BAT mit anderen Beat Alignment Tests, welche in verschiedenen Testparadigmen zur Puls-wahrnehmung validiert werden konnten (Dalla Bella et al. 2017; Fujii und Schlaug 2013; Iversen und Patel 2008; Müllensiefen et al. 2014; Ross et al. 2018). Jedes Testitem des Computerised Adaptive Beat Alignment Tests besteht aus einem kurzen Musikausschnitt und einer darüber gelegten Metronomspur. Die Aufgabe der Versuchspersonen ist es anzugeben, ob das Metronom und der implizite Puls des Musikstückes absolut synchron sind oder nicht. Dabei konstituiert sich die Item-Schwierigkeit hauptsächlich über den Abstand zwischen Metronom und musikalischem Puls. Der Test zeigt mit 25 Items eine annehmbare Test-Retest-Reliabilität,  $r(50) = 0,67$ ,  $95\%CI = [0,48; 0,80]$ , und korreliert mit der Subskala „musikalische Training“ des Gold-MSI Selbstauskunftfragebogens,  $r(69) = 0,41$ ,  $p < 0,001$ ,  $95\%CI = [0,20; 0,59]$  (Harrison und Müllensiefen 2018). Für die vorliegende Stichprobe liegt der *MSE* bei 0,95,  $95\%CI = [0,94; 0,96]$ .

**Melodic Discrimination Test (MDT)** Der MDT (Harrison et al. 2017) misst die Fähigkeit, Melodien im Arbeitsgedächtnis zu speichern und miteinander zu ver-

gleichen. In einem forced-choice Design werden drei verschiedene Versionen einer Melodie nacheinander präsentiert, wobei die Versuchspersonen angeben müssen, welche Version sich von den beiden anderen unterscheidet. Die Itemschwierigkeit wird über Parameter Melodielänge, Tonalität und Kontur modelliert. In der vorliegenden Stichprobe liegt der  $MSE$  bei 0,86,  $95\%CI = [0, 85; 0, 88]$ . Beim MPT kommt es für die Test-Retest-Reliabilität schon nach 11 Items zu einem Plateau bei einem Wert von  $r = 0,67$  und die Korrelation mit „musikalischem Training“ nach 10 Items ein Plateau von  $r = 0,44$  (Harrison et al. 2017).

**Mistuning Perception Test (MPT)** Im MPT (Larrouy-Maestri et al., 2019) müssen die Teilnehmenden heraushören, ob Gesang und Begleitung eines kurzen Musikausschnitts zueinander verstimmt sind oder nicht, also ob die Sängerin oder der Sänger richtig intoniert oder nicht. Dabei wird die Item-Schwierigkeit hauptsächlich über den Abstand von Gesangsspur zu den anderen musikalischen Spuren im Frequenzbereich bestimmt. In der vorliegenden Stichprobe liegt der  $MSE$  bei 0,76,  $95\%CI = [0, 75; 0, 77]$ . Für diesen Test wurde eine Test-Retest-Reliabilität von  $r = 0, 58$ ,  $95\%CI = [0, 39; 0, 72]$  bei 15 und von  $r = 0, 70$ ,  $95\%CI = [0, 54; 0, 81]$  bei den maximalen 30 Items gemessen. Auch hier stabilisiert sich die Korrelation mit der Subskala „musikalisches Training“ nach wenigen Items und liegt nach 10 Items um die  $r = 0,4$ .

### 3.3.2 Unabhängige Variablen

#### **Goldsmith Musical Sophistication Index (Gold-MSI, Müllensiefen et al. 2014)**

Dieser Selbstauskunftsfragebogen wurde genutzt, um die musikalische Expertise und Erfahrungen der Teilnehmenden mithilfe von fünf Subskalen zu erfassen. Für diese Studie ist die Subskala „Musikalische Ausbildung“ von Bedeutung, die aus sieben Fragebogenitems besteht, die auf einer 7-stufigen Skala beantwortet werden. Die Subskala misst das Ausmaß musikalischen Trainings in der Vergangenheit mithilfe von Fragen nach Trainingsdauer, -umfang und -erfolg. Die deutsche Übersetzung (Schaal et al. 2014) repliziert die Faktorstruktur des weit verbreiteten englischen Fragebogens und wurde zusätzlich an einer Stichprobe mit Schülerinnen und Schülern an weiterführenden Schulen validiert (Fiedler und Müllensiefen 2015). Die interne Konsistenz der englischen Subskala liegt bei  $\alpha = 0,903$  (deutsche Subskala:  $\alpha = 0,88$ ) und die Test-Retest-Korrelation beträgt  $r = 0,974$  (Müllensiefen et al. 2014).

**Concurrent musical activities (CCM, Müllensiefen et al. 2015)** Mit Hilfe des CCM Fragebogens werden die musikalischen Aktivitäten während der letzten drei Monate vor dem Erhebungszeitpunkt erfasst. Der CCM ist damit komplementär zur Subskala Musikalisches Training des Gold-MSI, welcher die musikalische Aktivität über längere Zeiträume erfasst. Wie in Müllensiefen et al. (2015) erläutert, erfasst der CCM über eine Checkliste, welche musikalische Aktivitäten in den letzten drei Monaten ausgeübt wurden (z.B. Spiel in Orchester/Ensemble, Musikmachen mit Freunden, Einzel- oder Gruppenunterricht etc.). Die binären Daten der Checkliste werden durch ein Raschmodell zu einem numerischen Wert zusammengefasst. Zusätzlich wird über zwei 7-Punkt-Skalen nach dem Ausmaß musikalischer Aktivität

und musikalischen Trainings in den letzten drei Monaten gefragt. Ein Hauptkomponentenmodell aggregiert die drei numerischen Variablen des CCM zu einem einzigen Wert und erklärt 75 % der Varianz der Variablen.

**Musical Home Environment (MHE, Ruth und Müllensiefen 2021)** Mithilfe der MHE Skala wird die musikalische Aktivität anderer Personen im selben Haushalt sowie die Unterstützung der eigenen musikalischer Aktivitäten durch die Eltern auf insgesamt vier binären Items und zwei Ratingskalen erfasst (Ruth und Müllensiefen 2021). Ähnlich wie beim CCM werden die binären Daten mit einem Raschmodell zusammengefasst dessen Output dann mit den beiden numerischen Variablen in einem Hauptkomponentenmodell aggregiert wird. Alle Variablen weisen Ladungen von  $>0,5$  auf der einzigen Komponente des Modells auf, was 67 % der Varianz der drei Variablen erklärt.

**Theory of Intelligence and Goal Choice (TOI)** Der TOI Fragebogen (Dweck 2000; Dweck et al. 1995) ist ein weit verbreitetes Selbstauskunftsinventar und misst das sogenannte „growth mindset“ der Versuchspersonen über die Einstellung zur Veränderbarkeit der eigenen Intelligenz und die Wahl von selbst gesetzten Zielen. Die Fragen der beiden Subskalen (Theory of Intelligence und Goal Choice) werden zusammen in einem Block präsentiert, aber getrennt ausgewertet. Sechs Fragen verwenden eine 6-Punkt Likert-Skala, die siebte Frage hat in binäres Antwortformat.

**Theory of Musicality (TOM, Eisinger 2021)** Der TOM Fragebogen erfasst die subjektiven Theorien über Musikalität und ist vom TOI-Fragebogen von Dweck inspiriert. Im Gegensatz zum TOI verwendet der TOM eine hierarchische Skalenstruktur, die an Biddle et al. (2003) angelehnt ist. Mit dem TOM Fragebogen wird erhoben, ob die Teilnehmenden musikalische Fähigkeiten eher als feststehende Begabung oder als Fähigkeit ansehen, die durch Training veränderbar ist. Insgesamt umfasst der TOM zwölf Fragen zur subjektiven Einschätzung von Veränderbarkeit und dem Ursprung von Musikalität. Die in der vorliegenden Untersuchung genutzten Faktoren erster Ordnung (Subskalen „entity“ & „incremental“, siehe Abb. 6 in Eisinger 2021) weisen einen Cronbachs-alpha von 0,63 und 0,79 auf.

**Ten Item Personality Inventory (TIPI, Gosling et al. 2003)** Das TIPI ist ein kurzes Persönlichkeitsinventar auf Grundlage des Big Five-Persönlichkeitsmodells (Goldberg 1990), das für die Verwendung mit Kindern und Jugendlichen leicht adaptiert wurde (Müllensiefen et al. 2015). Der TIPI wurde hinsichtlich Validität der Subskalen optimiert, die leicht besser ist als die anderer kurzer Persönlichkeitsinventare (Furnham et al. 2008). Die Test-Retest Reliabilität der fünf Subskalen des TIPI liegt zwischen 0,62 und 0,77 (Gosling et al. 2003) und damit größtenteils in einem akzeptablen Bereich.

**Intelligenz (MyIQ Test, Chan und Kosinski 2015; Condon und Revelle 2014)** Der MyIQ-Test ist eine adaptive und IRT-basierte Version der Progressiven Ravens Matrizen, die vom Cambridge Psychometric Centre als offener Test für allgemeine Intelligenz und fluide Intelligenz entwickelt wurde. Es ist allerdings zu

beachten, dass die Messung fluider Intelligenz als Fähigkeitskonstrukt i. d. R. den Einsatz mehrerer Tests für unterschiedliche inhaltliche Bereiche und verschiedene Aufgabenparadigmen erfordert. Der MyIQ-Test kann daher nur als Proxy für fluide Intelligenz gelten. Der *MSE* dieses Tests liegt in der vorliegenden Untersuchung bei 0,88, 95%CI = [0, 88; 0, 89].

**Jack and Jill Working Memory Test (JAJ, Tsigeman et al. 2022)** Durch den JAJ Test wird das räumlich-visuelle Arbeitsgedächtnis getestet. Der JAJ ist als adaptiver IRT-Test implementiert, orientiert sich am Arbeitsgedächtnismodell von Baddeley und Hitch (1974) und basiert auf einem Dual-Task-Paradigma. Die Länge der zu merkenden Sequenzen bestimmt dabei die Schwierigkeit der präsentierten Items. In der vorliegenden Stichprobe der LongGold-Studie liegt der *MSE* dieses Tests bei 0,36, 95%CI = [0, 36; 0, 37]. Mit JAJ-Test wurde bewusst ein Arbeitsgedächtnistest gewählt, bei dem keine auditorischen oder verbalen Informationen verarbeitet werden müssen. Das bedeutet, dass die Testergebnisse unabhängig von den Subsystemen des Arbeitsgedächtnisses sind (z. B. phonologische Schleife), von denen angenommen wird (z. B. Berz 1995), dass sie in einem engen Zusammenhang mit der Verarbeitung von musikalischen Informationen stehen. Das bedeutet, dass das visuell-räumlich Arbeitsgedächtnis als Prädiktor interpretiert werden kann, der seinerseits wahrscheinlich nur in einem geringen Maße (wenn überhaupt) durch musikalische Aktivität beeinflusst wird (s. a. Sala und Gobet 2017; Silas et al. 2022).

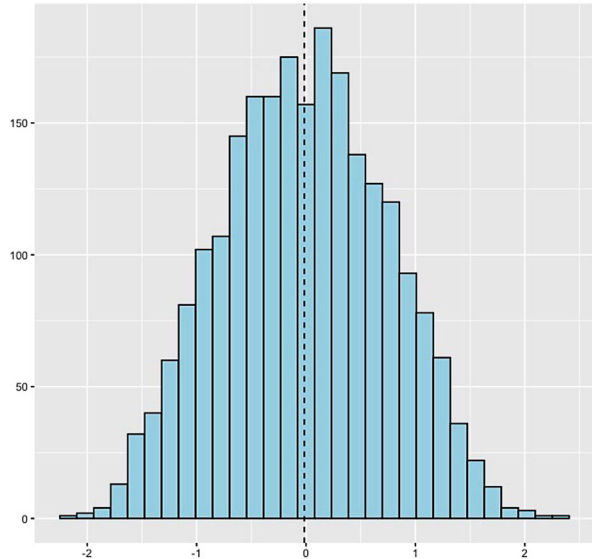
**Demographie und akademische Leistungen** Alter, Geschlecht, Nationalität sowie weitere Merkmale werden über einen demographischen Fragebogen von den Teilnehmenden erhoben. Zeugnisnoten aller Teilnehmenden werden anonymisiert von den Schulen übermittelt und als standardisierte Werte pro schulischem Fachbereich (Sprachen, MINT, Sozial- und Kulturwissenschaft, künstlerische Fächer, angewandte Fächer) sowie als Gesamtwert erfasst.

## 4 Analyse und Ergebnisse

### 4.1 Messung musikalischer Leistungsfähigkeit unabhängig vom Instrumentalspiel und ohne Voraussetzung formaler musikalischer Vorkenntnisse oder Terminologie

Dem Ansatz von Pausch et al. (2022) folgend wurde ein Generalfaktor (G-Faktor) aus den Testscores der drei musikalischen Hörtests (CA-BAT, MPT und MDT) mithilfe einer explorativen Minimum-Residual-Faktorenanalyse anhand der Funktion  $fa()$  des R-Pakets „psych“ (Revelle 2021) berechnet. Für die Analyse wurden nur Fälle verwendet, die den ersten Erhebungszeitpunkt eines/einer Teilnehmenden darstellen und aus den Klassen 5 oder 6 stammen. So können die Schülerinnen und Schüler (SuS) identifiziert werden, die bei ihrem Eintritt in die Studie musikalisch leistungsstark sind, ohne dass das Eintrittsalter als konfundierender Faktor wirkt. Insgesamt wurden so die Daten von  $N = 3090$  SuS für die Faktorenanalyse verwendet. Das resultierende Faktormodell erklärte 56% der Varianz der Testscores der

**Abb. 1** Histogramm mit G-Faktorwerten der Grundgesamtheit und eingezeichnetem Mittelwert



drei Hörtests, d. h. die Daten legen nahe, dass eine latente Variable einen substantiellen Anteil der Varianz der Testdaten erklärt und für die Korrelationen der drei Tests verantwortlich ist. Diese latente Variable wird als perzeptuelle musikalische Leistungsfähigkeit interpretiert und durch aggregierte Scores eines Faktormodells (G-Faktor) gemessen. Die Verteilung des so ermittelten Wertes des G-Faktors weist keine besonderen Ausreißer oder Schiefe auf (siehe Abb. 1). Die Messung des Konstrukts der perzeptuellen musikalischen Leistungsfähigkeit setzt somit weder voraus, dass die Versuchspersonen ein Instrument beherrschen, noch sonst über irgendwelche speziellen musikalischen Vorkenntnisse oder Terminologie verfügen.

#### 4.2 Identifikation musikalisch Begabter anhand von statischen Leistungskriterien

In Anlehnung an den z. B. von Gagné (1998) vertretenen Ansatz, Begabung anhand von überdurchschnittlicher musikalischer Leistung festzustellen, wurden zwei verschiedene Methoden verwendet, um über insgesamt vier Schwellenwerte leistungsstarke SuS zu identifizieren. Die erste Methode verwendet die Verteilung der Personenscores auf dem G-Faktor. Gemäß den Kriterien von Gagné (1998) wurden diejenigen SuS als musikalisch leistungsstark eingestuft, deren G-Faktor Wert den obersten 1% und 10% in der Verteilung des G-Faktors entsprechen.

Die zweite Methode berücksichtigt, dass musikalisches Training die Hörfähigkeiten verbessert, die dem G-Faktor zugrunde liegen. Um das Fähigkeitsniveau unabhängig von dem genossenen musikalischen Training zu ermitteln, orientiert sich die zweite Methode an modellbasierten Residuen, d. h. an dem Anteil der musikalischen Leistung, der nicht durch das erhaltene Training erklärt werden kann. Um auch nicht-lineare Zusammenhänge zu berücksichtigen, wurden Generalized Additive Models (GAMs, Wood 2017) zur Modellierung benutzt. Als abhängige Variable



wurden die G-Faktor-Werte verwendet und als Prädiktoren mehrere Variablen, die unterschiedliche Aspekte von musikalischem Training und Förderung messen (i.e. Subskala Musikalische Ausbildung des Gold-MSI, CCM, MHE). Als musikalisch besonders leistungsstark wurden die obersten 1% und 10% der Residualverteilung der GAM-Modelle definiert, d. h. diejenigen SuS die eine musikalische Leistung hatten, die deutlich höher war als anhand von musikalischem Training und Förderung im häuslichen Umfeld zu erwarten war.

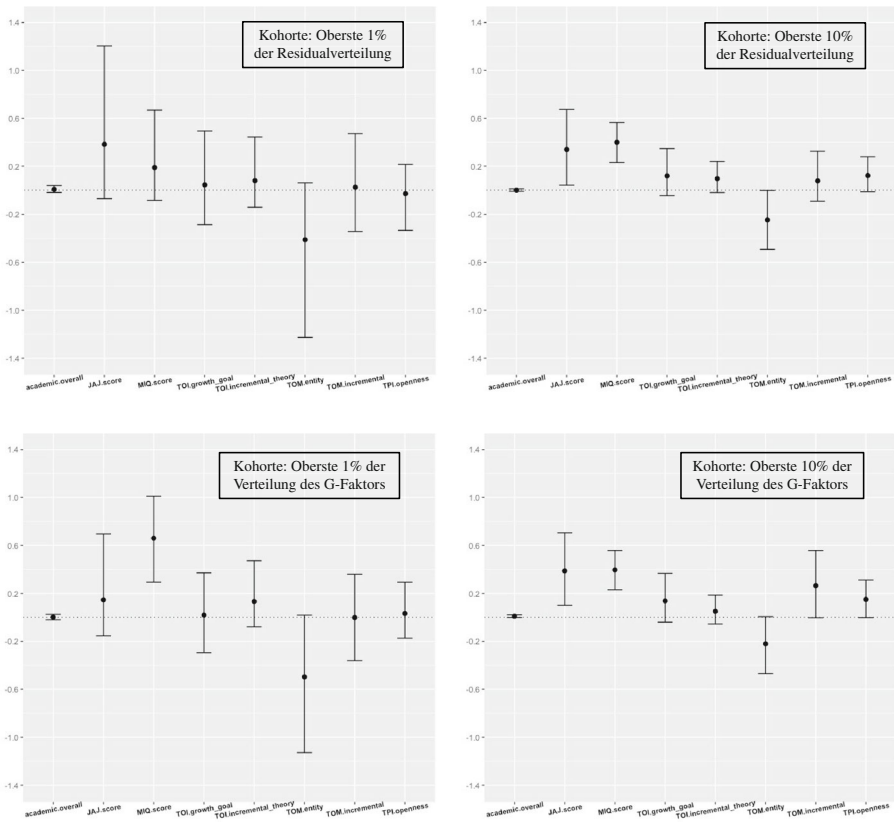
Insgesamt wurden durch die Kombination der beiden Methoden und beider Schwellenwerte vier neue binäre Variablen gebildet, die die SuS des Datensatzes jeweils als „begabt“ und „nicht-begabt“ im Sinne der Verteilungskriterien einteilen. Für die obersten 10% und 1% der Verteilung des G-Faktors (d. h. die Gruppe der *Begabten*) waren dies  $N=218$  und  $N=36$  SuS. Die obersten 10% und 1% der Residualverteilung (d. h. die *Begabten*) der GAM-Modelle entsprachen  $N=219$  und  $N=22$  SuS. Der Anteil der Schnittmenge der beiden Kriterien war bezogen auf das Leistungskriterium der obersten 10% der G-Faktor Verteilung 75,2% und für die obersten 10% des Residualkriteriums 74,9%.

### 4.3 Extra-musikalische Variablen als Prädiktoren für musikalische Begabte, die anhand von statischen Leistungskriterien ermittelt wurden

Um jene extra-musikalischen Variablen zu bestimmen, die mit dem musikalischen Leistungsniveau der SuS substantiell und robust assoziiert sind, wurden vier Bayesische logistische Regressionsmodelle für die vier verschiedenen Kriterien berechnet<sup>4</sup>. Als abhängige Variable diente jeweils das in 4.2 ermittelte binäre Kriterium der musikalischen Begabung und als mögliche Prädiktoren die in Abschn. 3.1.2 aufgeführten unabhängigen Variablen (JAJ, MyIQ, TOI, TOM, akademische Gesamtleistungen, TIPI-Faktor „Openness“). Zur Variablenselektion wurden sogenannte „horseshoe priors“ verwendet, die Regressionskoeffizienten mit geringer Effektstärke gegen null schrumpfen, um praktisch nicht bedeutsame Effekte zu eliminieren (van Erp 2020). Die Ergebnisse wurden mithilfe des R-Pakets „brms“ (Bürkner 2018) berechnet und sind in Abb. 2 zusammengefasst.

Wie zu erwarten, sind die kritischen Intervalle der Regressionskoeffizienten für die größere Stichprobe (d. h. obersten 10%) kleiner, weil die empirische Basis für die Schätzung größer ist. Weitergehend zeigt sich, dass die Tests zu Arbeitsgedächtnis und Intelligenz die höchsten Koeffizienten aufweisen und positiv mit dem musikalischen Begabungsstatus zusammenhängen. Als einzige Variable besitzt TOM.entity einen negativen Koeffizienten, der darauf hinweist, dass SuS, die Musikalität als unveränderliche Gabe auffassen, durch geringere musikalische Leistung gekennzeichnet sind. Zusätzlich weist auch der Persönlichkeitsfaktor „Offenheit“ des TIPI Persönlichkeitsinventars eine substantielle positive Assoziation mit dem Begabungsstatus auf.

<sup>4</sup> Die teilweise fehlenden Werte auf einigen Variablen wurden bei den nachfolgenden Analysen mithilfe einer multiplen Imputation kompensiert (Methode: predictive mean matching,  $m=50$  imputierte Datensätze, R-Paket „mice“, van Buuren und Groothuis-Oudshoorn 2011).



**Abb. 2** Punktdiagramme der Bayesschen logistischen Regressionsmodelle mit Konfidenzintervallen. Anmerkung: Darstellung zeigt die logistischen Regressionskoeffizienten mit den jeweiligen 95 % Bayesschen Konfidenzintervallen als Balken (Hespanhol et al. 2019)

#### 4.4 Entwicklung perzeptueller musikalischer Fähigkeiten bei Begabten, die anhand von statischen Leistungskriterien ermittelt wurden

Für die Modellierung der Entwicklung musikalischer Fähigkeiten wurden nur die Daten der SuS verwendet, von denen mindestens drei Erhebungszeitpunkte verfügbar waren. Deskriptive Daten dieser Stichprobe befinden sich in Tab. 6, 7 und 8 im Anhang. Mit Bayesschen linearen Mixed-Effects-Modells wurde anschließend die Entwicklung der perzeptuellen musikalischen Fähigkeiten (d. h. der G-Faktor-Werte) über die Zeit hinweg untersucht. Dabei wurden die unterschiedlichen Leistungsniveaus der SuS über modellbasierte „random intercepts“ (d. h. die individuellen Abweichungen vom mittleren Leistungsniveau) modelliert, sowie unterschiedliche Entwicklungsverläufe über sogenannte „random slopes“ (d. h. individuelle Abweichungen vom mittleren Entwicklungsverlauf). Abhängige Variable waren die G-Faktor-Werte und unabhängige Variablen waren die in der vorherigen Analyse (4.3) ermittelten Prädiktoren (JAJ, MyIQ, TOM.entity, TIPI.openness) und die Variable „Alter“, die den Entwicklungsverlauf über die Zeit modelliert.

Aufgrund fehlender Datenpunkte konnten für die 1 %-Perzentilgruppen des Verteilungs- und des Residualkriteriums keine Modelle berechnet werden. Die folgenden Ergebnisse basieren deshalb nur auf den beiden obersten 10 %-Kriterien. Es wurden separate Modelle für die SuS mithilfe beider Kriterien für die begabte Gruppe gerechnet und die nicht-Begabten in einem Modell zusammengefasst. Für die Operationalisierung von nicht-Begabten bedeutet dies, dass Teilnehmende, die nach nur einem der beiden statischen Begabungskriterien als begabt eingestuft wurden, in der Aufbereitung des Datensatzes der nicht-Begabten herausgefiltert wurden (vgl. Tab. 1).

Wie zu erwarten, besitzt die Gruppe der nicht-Begabten ein niedrigeres fixed effects intercept als die Gruppe der Begabten. Bemerkenswert ist dagegen die Abnahme der musikalischen Hörleistungen (der G-Faktor-Werte) bei den Begabten über die Zeit und das Wachstum der Hörleistungen bei den nicht-Begabten. Als Prädiktor ist Intelligenz (MyIQ) durchgehend positiv mit der Entwicklung der Hörleistungen

**Tab. 1** Modellkoeffizienten der drei Baysschen longitudinalen linear mixed model Analysen nach klassischen Begabungskriterien

Fixed Effects	<i>b</i>	<i>SE</i>	<i>CI<sup>a</sup></i>
<i>Oberste 10% G-Faktor; N = 50 Datenpunkte von 16 Teilnehmenden</i>			
Intercept	3,49	1,22	[1,18; 6,02]
JAJ	-0,02	0,04	[-0,09; 0,05]
MyIQ	0,08	0,04	[0,00; 0,15]
Alter	-0,16	0,08	[-0,33; -0,01]
TOM.entity	-0,14	0,12	[-0,39; 0,10]
TIPL.openness	0,02	0,07	[-0,12; 0,17]
<i>Oberste 10% Residuum, N = 71 Datenpunkte von 23 Teilnehmenden</i>			
Intercept	1,68	0,93	[-0,17; 3,48]
JAJ	-0,06	0,05	[-0,16; 0,04]
MyIQ	0,10	0,05	[0,01; 0,19]
Alter	-0,12	0,07	[-0,26; -0,02]
TOM.entity	0,07	0,13	[-0,19; 0,32]
TIPL.openness	0,09	0,07	[-0,05; 0,24]
<i>Nicht Begabte, N = 1840 Datenpunkte von 546 Teilnehmenden</i>			
Intercept	-1,10	0,23	[-1,55; -0,66]
JAJ	0,05	0,04	[-0,02; 0,12]
MyIQ	0,10	0,03	[0,04; 0,17]
Alter	0,09	0,01	[0,07; 0,12]
TOM.entity	-0,11	0,03	[-0,17; -0,06]
TIPL.openness	0,06	0,02	[0,02; 0,10]

Alle drei Modelle wurden mit  $N = 10.000$  iterations, 8 chains und 1000 warm-up samples mithilfe eines Hamiltonian Monte Carlo Algorithmus (Homan und Gelman 2014) geschätzt. Es wurden gaussische prior distributions und delta bei einem Wert von 0,99 festgelegt. Imputationen für die Variablen JAJ und MIQ wurden aus Gründen der besseren Modell- und Rechenleistung innerhalb des Modells mithilfe einer Regressionsgleichung imputiert (Bürkner 2021). Durch sehr gute Rhat-Werte (1,00–1,01) und Tail- sowie Bulk effective sample size Werte ( $> 1000$ ) ist anzunehmen, dass die Schätzung der a-posteriori Wahrscheinlichkeiten verlässlich ist

<sup>a</sup>95 % Bayessches credible interval (Hespanhol et al. 2019)

assoziiert. Durchweg positiv mit dem Entwicklungsverlauf der musikalischen Fähigkeiten assoziiert ist auch die Persönlichkeitsdimension „Openness“. Arbeitsgedächtnis (JAJ) und die subjektive Theorie der Musikalität zeigen dagegen uneinheitliche Assoziation und haben Koeffizienten mit breiten kritischen Intervallen, was allenfalls geringe oder Nulleffekte dieser Prädiktoren nahelegt. Die Modell-Fits sind für das Perzentilkriterium mit  $R_{\text{bayes}}^2 = 0,55$ , für das Residualkriterium mit  $R_{\text{bayes}}^2 = 0,55$  und die nicht-Begabten  $R_{\text{bayes}}^2 = 0,67$ , relativ hoch.

#### 4.5 Identifikation von Begabten nach dem TAD-Ansatz und unter Verwendung von Langzeitdaten

Der TAD-Ansatz verknüpft Begabung mit akzelerierten positiven Entwicklungsverläufen. Gleichzeitig wird Begabung als nicht direkt messbare, sondern latente Größe angenommen, die aber über geeignete Proxy-Variablen angenähert werden kann. Um diese Annahmen zu operationalisieren, wird ein statistisches Modell benötigt, was Begabung als latente Variable repräsentiert und gleichzeitig am Zuwachs der perceptuellen musikalischen Fähigkeiten über die Zeit festmacht. Soll die Einteilung des Begabungsstatus wie beim traditionellen Ansatz in zwei Klassen (begabt/nichtbegabt) erfolgen, dann bietet sich hierfür eine longitudinale Latent Class Analysis (LCA, Lin et al. 2002) an, welche mithilfe des R-Pakets „lcm“ (Proust-Lima et al. 2017) berechnet werden kann. In diesem Analysemodell wird davon ausgegangen, dass sich die Stichprobe in unterschiedliche Gruppen einteilen lässt, die sich in Bezug auf die Zunahme der Hörleistungen (G-Faktor) über die Zeit unterscheiden, d. h. unterschiedliche Gruppen von SuS können ggf. mithilfe der LCA anhand ihrer individuellen Wachstumskoeffizienten („random slopes“) identifiziert werden. Es wurden bei dieser Analyse nur diejenigen SuS berücksichtigt, die an drei oder mehr Erhebungszeitpunkten teilgenommen hatten, sodass ein Datensatz mit 1120 Datenpunkten von  $N = 364$  SuS die Grundlage für die nachfolgende Analyse bildet.

Um die Vergleichbarkeit mit der 2-Klasseneinteilung der statischen Begabungskriterien zu bewahren, wurden zwei Modelle mit jeweils zwei Klassen und unterschiedlichen „random effects“ Spezifikationen in der Modellauswahl betrachtet. Zur besseren Vergleichbarkeit und Einordnung der Fit-Werte, wurde zusätzlich ein Basismodell zum Abgleich herangezogen, bei dem die SuS nicht in unterschiedliche Klassen eingeteilt werden (= 1-Klassenmodell).

Nach der Betrachtung der Tab. 4 und 5 (siehe Anhang) konnte kein bedeutsamer Unterschied zwischen den 1- und 2-Klassenmodell anhand des BICs, AICs, der log-likelihood und der mittleren Klassifikationswahrscheinlichkeiten festgestellt werden. Das lässt darauf schließen, dass die Wachstumsraten musikalischer Leistungsfähigkeit kontinuierlich verteilt sind und keine empirischen Hinweise für eine strikte Einteilung in zwei Klassen (begabt/unbegabt) vorliegt. Um diesem Umstand in der Analyse gerecht zu werden, wurde im Weiteren ein Latent Growth Model berechnet, das auf eine Einteilung in verschiedene latente Begabungsklassen verzichtet und stattdessen Begabung als kontinuierlich verteilte Variable annimmt.

#### 4.6 Latentes Wachstumsmodell und Analyse extra-musikalischer Variablen als Prädiktoren für musikalischer Begabung (TAD-Ansatz)

Um die Vergleichbarkeit zwischen klassischen Begabungskriterien und TAD-Ansatz zu gewährleisten, wurde der Datensatz gefiltert. Als Obergrenze wurde ein Alter von 14 Jahren zum ersten Teilnahmezeitpunkt festgelegt, was dem höchsten Alter eines Sechstklässlers im Datensatz entspricht. Als statistischer Modellierungsansatz wurde das Latent Curve bzw. Growth Model (LGM/LCM) gewählt (Jung und Wickrama 2008). Diese gehören zur Familie der Strukturgleichungsmodelle und erlauben die Untersuchung longitudinaler Daten mit Messwiederholung. Dabei wird angenommen, dass die individuellen Wachstumskurven der Teilnehmenden anhand eines latenten Wachstumskoeffizienten („slope“) und des mittleren Fähigkeitsniveaus („intercept“) erklärt werden können (Ghisletta et al. 2015). Dieser an Wachstumskoeffizienten orientierte statistische Ansatz mit der Möglichkeit, zeitabhängige und zeitunabhängige Prädiktoren im selben Modell zu verwenden, ist dafür geeignet, Begabung als kontinuierlich verteilte Variable zu analysieren. Datengrundlage sind  $N=1095$  Datenpunkte, gesammelt über  $N=3$  Erhebungszeitpunkte von insgesamt  $N=365$  Teilnehmenden. Umgesetzt wurde das Modell mithilfe der Programmiersprache R und dem Softwarepaket „lavaan“ (Rosseel 2012).

Modell 1 (siehe Tab. 2) enthält noch keine Prädiktoren der latenten Faktoren „slope“ und „intercept“. Somit ist die Grundstruktur des Wachstumsmodells einfacher zu interpretieren und es kann ein Vergleich des voll ausspezifizierten Modells mit Prädiktoren und dem Basismodell (Modell 1) angestellt werden kann. Die Fit-Kriterien des Modells,  $CFI = 1$ ,  $TLI = 1$ ,  $\chi^2(1,356) = 0,01$  mit  $p = 0,94$ ,  $RMSEA = 0$  mit  $95\%CI[0; 0,03]$ ,  $SRMR = 0,001$ , zeigen eine sehr gute Passung von Modell und empirischen Daten an. Im Hinblick auf die latenten Wachstumskoeffizienten weist die Kohorte im Schnitt einen positiven Entwicklungsverlauf musikalischer Begabung auf.

**Tab. 2** Modellkoeffizienten des Latent Growth Models (TAD-Ansatz) ohne Prädiktoren der latenten Wachstumsfaktoren

	<i>b</i>	<i>SE</i>	<i>CI</i>
<i>Intercepts</i>			
Intercept	-0,14	0,04	[-0,22; -0,07]
Slope	0,06	0,02	[0,02; 0,11]
<i>Variances</i>			
G-Factor_t0	0,25	0,06	[0,14; 0,37]
G-Factor_t1	0,35	0,03	[0,28; 0,41]
G-Factor_t2	0,34	0,06	[0,22; 0,45]
Intercept	0,30	0,06	[0,19; 0,42]
Slope	0,04	0,03	[-0,01; 0,10]

Modellkoeffizienten des Latent Growth Models (TAD-Ansatz) ohne Prädiktoren der latenten Wachstumsfaktoren. Der Parameter ‚intercept‘ bezieht sich auf den mittleren G-Faktorenwert der Teilnehmenden zum ersten Erhebungszeitpunkt (t0) und der ‚slope‘-Parameter ist die Schätzung der mittleren Wachstumsrate im G-faktor der Teilnehmenden. Durch die ‚variances‘ werden die Varianz des G-Faktors zum Zeitpunkt t und die Residualvarianzen des ‚intercepts‘ und der ‚slope‘ geschätzt

**Tab. 3** Modellkoeffizienten des voll spezifizierten Latent Growth Models (TAD-Ansatz)

	<i>b</i>	<i>SE</i>	<i>CI</i>
<i>Intercepts</i>			
Intercept	-0,04	0,40	[-0,81; 0,73]
Slope	0,04	0,25	[-0,46; 0,53]
<i>Variances</i>			
G-Factor_t0	0,26	0,05	[0,16; 0,37]
G-Factor_t1	0,34	0,03	[0,28; 0,41]
G-Factor_t2	0,35	0,06	[0,23; 0,46]
Intercept	0,22	0,05	[0,12; 0,32]
Slope	0,03	0,03	[-0,02; 0,09]
<i>Regressionskoeffizienten intercept</i>			
MIQ	0,11	0,04	[0,04; 0,18]
TOI.growth	-0,03	0,05	[-0,13; 0,06]
TOI.incremental	-0,01	0,03	[-0,08; 0,06]
TOM.entity	-0,13	0,05	[-0,24; -0,03]
TOM.incremental	0,03	0,06	[-0,09; 0,16]
TPI.openness	0,09	0,03	[0,02; 0,15]
JAJ	0,16	0,05	[0,07; 0,25]
<i>Regressionskoeffizienten slope</i>			
MIQ	0	0,02	[-0,05; 0,05]
TOI.growth	0,03	0,03	[-0,03; 0,09]
TOI.incremental	0,01	0,02	[-0,03; 0,06]
TOM.entity	0,02	0,04	[-0,05; 0,09]
TOM.incremental	-0,02	0,04	[-0,10; 0,06]
TPI.openness	-0,02	0,02	[-0,06; 0,02]
JAJ	-0,07	0,03	[-0,12; -0,01]

Der Parameter ‚intercept‘ bezieht sich auf den mittleren G-Faktorenwert der Teilnehmenden zum ersten Erhebungszeitpunkt (t0) und der ‚slope‘-Parameter ist die Schätzung der mittleren Wachstumsrate im G-faktor der Teilnehmenden. Durch die ‚variances‘ werden die Varianz des G-Faktors zum Zeitpunkt t und die Residualvarianzen des ‚intercepts‘ und der ‚slope‘ geschätzt. Alle Prädiktoren die als Kovariaten mit den latenten Wachstumskoeffizienten modelliert wurden, sind stellen die Werte des ersten Erhebungszeitpunkts da und sind damit TAD-Konform

Modell 2 (siehe Tab. 3) beinhaltet zusätzliche Prädiktoren, die als Regressoren in einem Modell den Zusammenhängen mit den latenten Variablen des „intercepts“ und der „slope“ modellieren sollen. Auch hier weisen die Fit-Indizes eine sehr gute Passung von Modell und Daten auf,  $CFI = 1$ ,  $TLI = 1$ ,  $\chi^2(8,356) = 4,08$  mit  $p = 0,85$ ,  $RMSEA = 0$  mit  $95\%CI[0; 0,04]$ ,  $SRMR = 0,01$ . Dabei zeigt sich, dass fluide Intelligenz, die Leistung des Arbeitsgedächtnis und der Persönlichkeitsfaktor „Offenheit“ ein hohes anfängliches musikalisches Leistungsniveau begünstigen und die Einstellung, die Musikalität als unveränderliche Gabe aufzufassen, den gegenteiligen Effekt hat. Als einzig statistisch relevanter Prädiktor für eine hohe Wachstumsrate des musikalischen G-Faktors weist das Modell das Arbeitsgedächtnis aus. Da dieser Regressor ein negatives Vorzeichen besitzt, bedeutet dies, dass je größer anfängliche Kapazität des Arbeitsgedächtnisses ist, desto niedriger die

Wachstumsrate ausfällt. Durch den Umstand, dass das Einstiegsalter (9 bis 14 Jahre) und damit die Kapazität des Arbeitsgedächtnis zu Beginn der Testung innerhalb der Stichprobe stark variieren, ist davon auszugehen, dass die negative Korrelation mit der Wachstumsrate durch eine positive Korrelation von Einstiegsalter und Arbeitsgedächtnisleistung entstehen. Berücksichtigt man diesen Umstand und die Annahme, dass Arbeitsgedächtnis und musikalische Begabung generell positiv korreliert sind, lässt sich die negative Korrelation von Wachstumsrate so deuten, dass jüngere Teilnehmende eine niedrigere Arbeitsgedächtnisleistung aufweisen und diese aber über die Zeit aufholen. In Tab. 9 im Anhang ist ein zusätzliches Modell aufgeführt, was diese Hypothese prüft, in dem ein zusätzlicher Interaktionsterm von Einstiegsalter  $\times$  Arbeitsgedächtnis in das Modell aufgenommen wird. Dieser Interaktionsterm trägt in der Tat substanziell zur Erklärung der Wachstumsraten und ist positiv mit Wachstum assoziiert. Das bedeutet, dass eine größere Arbeitsgedächtniskapazität zumindest bei älteren SuS mit dem Wachstum musikalischer Leistungsfähigkeit assoziiert ist.

#### 4.7 Vergleich der beiden Begabungsansätze

Die Variablenselektion im klassischen Begabungsansatz und die Signifikanzniveaus der Regressoren im LGM nach TAD ergaben, wenn auch unterschiedlich zu interpretierende, Prädiktoren für die Klassifizierung der Teilnehmenden in „begabt“ und „nicht begabt“.

Bei den Modellen zur Entwicklung musikalischer Fähigkeiten zeigt sich zunächst anhand der „intercepts“, dass die nach traditionellen Kriterien als begabt klassifizierten SuS höhere musikalische Fähigkeiten besitzen als die unbegabten SuS. Allerdings zeigen die Begabten eine deutlich negative Entwicklung des G-Faktors über die Zeit. Dies steht im deutlichen Kontrast zu der Begabungskonzeption nach dem TAD-Ansatz, wo eine positive akzelerierte Entwicklung des G-Faktors als grundlegender Mechanismus musikalischer Begabung angenommen wird. Dadurch, dass in der TAD-Modellierung der Entwicklung musikalischer Fähigkeiten keine Schwellenwerte vorgegeben werden, wird für die Verantwortlichen in der Förderung musikalisch begabter Kinder und Jugendlicher offengelassen, ab welchen Werten eine Person als „begabt“ eingestuft bzw. für das jeweilige Programm als geeignet erachtet wird. In Bezug auf die möglichen Proxyvariablen musikalischer Begabung zeigen sich in den Modellen mit JAJ, MIQ, TOM.entity und TPI.openness, in den Modellen beider Ansätze dieselben Variablen als Prädiktoren für die musikalische Begabung zu Beginn des Erhebungszeitraumes. Im Vergleich des longitudinalen Modells der anhand klassischer Kriterien eingeteilten Begabungsgruppen und der nach TAD ermittelten latenten Wachstumsrate musikalischer Begabung wird deutlich, dass fluide Intelligenz, Arbeitsgedächtnis und Offenheit mit hohen G-Faktor-Werte einhergehen. Trotzdem haben diese Proxyvariablen keine oder nur eine geringe Vorhersagekraft für die musikalische Begabung im Sinne einer akzelerierten positiven musikalischen Entwicklung. Einzig das Arbeitsgedächtnis scheint bei älteren SuS im TAD-Modell positiv mit musikalischer Entwicklung assoziiert zu sein.

## 5 Diskussion

Ziel der vorliegenden Untersuchung war es, mithilfe zweier verschiedener Begabungsansätze (traditionelle Leistungskriterien vs. TAD-Ansatz) die jugendlichen Teilnehmenden einer Langzeitstudie in musikalische Begabte und Nicht-Begabte einzuteilen. Zusätzlich wurden in Querschnittsanalysen robuste Prädiktoren ermittelt, die mit dem Begabungsstatus der Jugendlichen assoziiert sind. Anschließend wurden dann für die beiden Begabungsgruppen Längsschnittmodelle berechnet, um den jeweiligen Entwicklungsverlauf musikalischer Hörleistungen zu untersuchen und die Prädiktoren aus den Querschnittsanalysen aus der Perspektive der Längsschnittuntersuchungen zu validieren. Diese wurden mit einem Längsschnittmodell verglichen, das die individuellen Wachstumsraten berücksichtigt.

### 5.1 Unterschiede zwischen den Begabungsansätzen bei der Entwicklung perzeptueller musikalischer Leistungen

Als größter Unterschied zwischen den beiden Begabungsansätzen lässt sich der Entwicklungsverlauf innerhalb der Gruppe der Begabten festhalten. Für die nach traditionellen Leistungskriterien klassifizierten Begabten liefern das Längsschnittmodell zum musikalischen Entwicklungsverlauf kontraintuitive Ergebnisse. Hier lässt sich der Einfluss von Messfehlern und das Phänomen der Regression zur Mitte (vgl. Jerrim und Vignoles 2013) an den Modellen ablesen, die auf der Einteilung der SuS nach traditionellen Leistungskriterien beruhen. Die enorm hohen individuellen intercepts der Begabten zusammen mit der abwärts gerichteten Entwicklung der musikalischen Hörleistungen der SuS sind Indizien dafür, dass diese Entwicklungsmodelle nicht plausibel sind. Gerade weil es ein Alter betrifft, in dem sich die musikalischen Fähigkeiten gerade bei begabten SuS noch weiter entwickeln sollten. Es scheint also möglich, dass ein Anteil der hohen oder niedrigen G-Faktor-Werte der ersten Datenerhebung als messfehlerbedingte statistische Abweichungen zu betrachten sind, die sich im weiteren Verlauf der Studie in Richtung des Gruppenmittelwerts bewegen.

Für die Praxis der Identifikation von musikalisch begabten Kindern und Jugendlichen bedeutet dies, dass die Anwendung von traditionellen Leistungskriterien auf einem Querschnittsdatensatz möglicherweise dazu führt, dass Individuen mit hohem Leistungspotential nicht entdeckt werden. Die so als begabt eingestuften Individuen können im weiteren Zeitverlauf sogar eine geringere musikalische Leistung als die sich schneller entwickelnden Gleichaltrigen aufweisen, die zunächst als weniger begabt eingestuft wurden. Die Ergebnisse der nach dem TAD-Ansatz modellierten LGMs zeigen demgegenüber, dass erst nach wiederholtem Testen die Einstufung als begabt robuste Aussagen darüber erlaubt, welche Individuen neben überdurchschnittlichen musikalischen Leistungen auch ein hohes Entwicklungspotential haben.

### 5.2 Prädiktoren für musikalische Begabung

Gemeinsam ist den Ergebnissen aus beiden Ansätzen, dass vor allem die fluide, allgemeine Intelligenz ein relevanter Prädiktor für musikalische Begabung zum ersten Erhebungszeitpunkt darstellen. In Bezug auf den weiteren Verlauf musikalischer



Entwicklung zeigen sich im TAD-konformen LGM, abgesehen vom positiven Effekt des Arbeitsgedächtnisses bei älteren SuS noch keine robusten Ergebnisse, was seinen Grund unter anderem an den noch wenigen Erhebungszeitpunkten und durch diverse Filterungen verkleinerten Datensatz haben kann. Insgesamt konnten in der vorliegenden Studie die Leistungen des räumlich-visuellen Arbeitsgedächtnisses, der fluiden sowie allgemeinen Intelligenz, der Persönlichkeitsfaktor „Offenheit“ und die Überzeugung, dass Musikalität und Intelligenz veränderbare Konstrukte sind, als relevante Proxyvariablen identifiziert werden. Die hier vorgestellte Implementierung des TAD-Ansatzes erfüllt also einige der Forderungen, die McIntosh et al. (2018) an moderne Modelle zur Identifizierung von Hochbegabung stellen: Die Fixierung auf absolute Schwellenwerte wird vermieden und die Leistungsentwicklung über einen Zeitraum fließen in die Bewertung des Begabungsstaus mit ein. Zudem erlaubt die Operationalisierung als latentes Variablenmodell die Trennung des Messfehlers vom Zielkonstrukt Begabung. So können kontraintuitive Ergebnisse, wie die negative Entwicklung von Hörleistungen bei musikalisch Begabten, vermieden werden (Goldstein und French 2015). Allerdings bleibt es eine Aufgabe für zukünftige Studien mit robusteren Ergebnissen zu belegen, welche psychologischen Faktoren (d. h. Proxyvariablen) für die musikalische Entwicklung innerhalb des TAD-Ansatzes eine bedeutsame Rolle spielen.

### 5.3 Limitationen und Erweiterungen in zukünftigen Studien

Als limitierender Faktor dieser Studie ist zu nennen, dass für die meisten SuS nur drei Messzeitpunkte zur Verfügung standen. Da innerhalb der LongGold-Studie gegenwärtig weitere Daten gesammelt werden, sollte es aber in Zukunft leicht möglich sein, die hier erarbeiteten Modellen mit einer größeren Stichprobe, die weitere Messzeitpunkte für dieselben SuS enthält, zu überprüfen.

Eine weitere Einschränkung ist, dass die Stichprobe für die Untersuchung extremer Hochbegabung nicht groß genug bzw. nicht spezialisiert genug ist. Das bedeutet, dass die hier dargestellten Ergebnisse sich hinsichtlich herausragender oder extremer musikalischer Fähigkeiten oder Begabung nicht einfach generalisieren lassen. Eine Validierung der hier vorgestellten Modelle anhand von spezialisierten Populationen (z. B. SuS an Musikschulen oder allgemeinbildenden Schulen mit Musikschwerpunkt) wäre demnach ein wichtiger Schritt, um Aussagen über Entwicklungsverläufe und relevante Prädiktorvariablen auch bei musikalisch Hochbegabten und besonders Leistungsstarken zu ermöglichen.

In Bezug auf die Operationalisierung musikalischer Fähigkeiten ist festzuhalten, dass musikalische Hörtests nur eine von mehreren Fähigkeitsdimensionen abdecken und die Ergebnisse dieser Studie wahrscheinlich nicht eins zu eins auf andere musikalische Fähigkeiten, wie beispielsweise das Instrumentalspiel, übertragbar sind. Jedoch ist es möglich, den hier verwendeten musikalischen G-Faktor gegen ein latentes Konstrukt auszutauschen, das die Fähigkeit zur musikalischen Produktion repräsentiert, wenn entsprechende Indikatorvariablen erhoben werden können. Mit Ausblick auf zukünftige Untersuchungen zur Identifikation und Entwicklung musikalischer Begabung wäre es zudem wünschenswert, weitere potenzielle Prädiktoren für musikalische Begabung, bspw. den Durchhaltewillen (*grit*, Duckworth et al.

2007), zu berücksichtigen, die in der Literatur derzeit diskutiert werden. Insgesamt zeigt sich anhand der Nulleffekte bzw. im Rahmen theoretischer Annahmen schwer interpretierbarer Ergebnisse in Bezug auf die Proxyvariablen, dass hier noch weiterer Forschungsbedarf besteht. Vor allem mit dem Alter konfundierte Variablen, wie die Leistung des visuell-räumlichen Arbeitsgedächtnis, sollten in Zukunft anhand von Normwerten in die Analyse miteinbezogen werden, was aufgrund der aktuellen Datenlage für diese Untersuchung noch nicht realisierbar war. In Bezug auf die Erhebungszeiträume könnte es zusätzlich sinnvoll sein, die Entwicklung musikalischer Fähigkeiten in kürzeren Zeitintervallen (z. B. 3 oder 6 Monate) zu erheben, sowie auch produktive musikalische Fähigkeiten (z. B. Instrumentalspiel) zu messen.

#### 5.4 Bedeutung der Ergebnisse und Schlussfolgerungen

Trotz der genannten Einschränkungen besteht ein Vorteil dieser Studie in der Verwendung von allgemein verständlichen Hörtests und einer breiten, nicht spezialisierter Stichprobe. Dies erlaubt es, die Ergebnisse im Sinne einer breiten Population zu interpretieren, für die weder Instrumentalkenntnisse, intensiver Musikunterricht noch besondere Kenntnisse in Musiktheorie oder -terminologie vorausgesetzt werden müssen. Das bedeutet, dass die hier verwendeten Methoden es möglich machen, musikalische Begabung auch für Individuen zu abschätzen, die aufgrund sozialer oder ökonomischer Nachteile keine Instrumentalerfahrung sammeln oder außerschulischen musikalischen Aktivitäten nachgehen konnten. So können nach dem TAD-Ansatz Individuen, die zunächst nicht durch musikalische Höchstleistungen auffallen, aber trotzdem ein überdurchschnittliches musikalisches Entwicklungspotential besitzen, als begabt identifiziert und für musikalische Förderung ausgewählt werden.

Die Kombination aus Tests, die keine besonderen musikbezogenen Voraussetzungen besitzen, und der hier vorgeschlagenen statistischen Operationalisierung des TAD-Ansatzes eröffnet also neue Möglichkeiten für die Identifikation und Entwicklung von musikalischer Begabung. Zusammen mit der Nutzung von Stimulusmaterialien, die sich nicht auf die westliche Kunstmusik beschränken, kann dieser Ansatz somit ein Schritt in Richtung einer inklusiveren Definition von musikalischer Begabung sein, die Musik in ihrer stilistischen und praktischen Breite in der Gesellschaft repräsentiert.

## 6 Anhang

Nach dem ersten Schritt der Modellselektion (siehe Tab. 4) zeigt sich, dass AIC und die log-likelihood der Modelle ähnliche Fit-Werte aufweisen. Für die weiteren Berechnungen wurde sich daher für die Spezifikation von Modell drei anhand des BICs entschieden (2 latente Gruppen, gemeinsamer intercept im mixture model), da der gemeinsame intercept der beiden Klassen eher den Überlegungen des TAD-Ansatzes entspricht.

**Tab. 4** Modellfitkoeffizienten der verschiedenen Zielmodellkandidaten der LCA

Mixture Modell	Gruppen	loglik	BIC	AIC
1. Ohne mixture Modell	1	-1160	2355	2332
2. (1 Alter)	2	-1154	2461	2326
3. -1 + (1 Alter)	2	-1155	2358	2327

**Tab. 5** Mittlere Klassifikationswahrscheinlichkeiten innerhalb der durch LCA ermittelten Klassen

Mixture Modell		Nicht-Begabt	Begabt
2. (1 Alter)	Nicht-Begabt	0,86	0,14
	Begabt	0,18	0,82
3. -1 + (1 Alter)	Nicht-Begabt	0,88	0,12
	Begabt	0,21	0,79

**Tab. 6** Alterswerte aller in der Analyse verwendeten Erhebungszeitpunkte

		Alter					
		DE			UK		
		<i>n</i>	<i>m</i>	<i>sd</i>	<i>n</i>	<i>m</i>	<i>sd</i>
Testzeitpunkt	1	235	10,4	0,64	247	13	1,05
	2	235	11,4	0,65	235	14	1,05
	3	235	12,4	0,98	212	14,9	1,00
	4	-	-	-	124	16,2	1,03
	5	-	-	-	52	16,9	0,75

**Tab. 7** Akademische Leistungen aller in der Analyse verwendeten Erhebungszeitpunkte

		Akademische Leistungen (alle Fächer)					
		DE			UK		
		<i>n</i>	<i>m</i>	<i>sd</i>	<i>n</i>	<i>m</i>	<i>sd</i>
Testzeitpunkt	1	234	76,9	14,5	322	69	12,1
	2	235	71,1	15,4	310	69,1	12,0
	3	235	67,2	16,9	276	66,4	12,0
	4	-	-	-	171	64,3	12,9
	5	-	-	-	66	62,1	13,2

**Tab. 8** My.IQ Werte aller in der Analyse verwendeten Erhebungszeitpunkte

		MIQ					
		DE			UK		
		<i>n</i>	<i>m</i>	<i>sd</i>	<i>n</i>	<i>m</i>	<i>sd</i>
Testzeitpunkt	1	235	-2,23	0,93	230	-1,33	1,01
	2	205	-2,06	1,04	226	-1,21	1,04
	3	198	-1,94	1,07	253	-1,16	1,08
	4	-	-	-	124	-1,07	1,14
	5	-	-	-	55	-1,15	1,43

**Tab. 9** Modellkoeffizienten des voll spezifizierten Latent Growth Models mit Interaktion

	<i>b</i>	<i>SE</i>	<i>CI</i>
<i>Intercepts</i>			
Intercept	-0,42	0,41	[-1,22; 0,39]
Slope	-0,35	0,28	[-0,90; 0,21]
<i>Variances</i>			
G-Factor_t0	0,26	0,05	[0,15; 0,36]
G-Factor_t1	0,35	0,03	[0,29; 0,41]
G-Factor_t2	0,33	0,06	[0,22; 0,44]
Intercept	0,22	0,05	[0,12; 0,32]
Slope	0,04	0,03	[-0,01; 0,09]
<i>Regressionskoeffizienten intercept</i>			
MIQ	0,10	0,04	[0,03; 0,17]
TOI.growth	-0,03	0,05	[-0,13; 0,07]
TOI.incremental	-0,01	0,03	[-0,08; 0,05]
TOM.entity	-0,14	0,05	[-0,25; -0,04]
TOM.incremental	0,03	0,06	[-0,09; 0,15]
TPI.openness	0,08	0,03	[0,02; 0,15]
IAJ	0,16	0,04	[0,07; 0,25]
Einstiegsalter * IAJ	0,04	0,01	[0,01; 0,06]
<i>Regressionskoeffizienten slope</i>			
MIQ	-0,01	0,02	[-0,06; 0,04]
TOI.growth	0,03	0,03	[-0,03; 0,10]
TOI.incremental	0,01	0,02	[-0,03; 0,05]
TOM.entity	0,02	0,04	[-0,05; 0,08]
TOM.incremental	-0,02	0,04	[-0,10; 0,06]
TPI.openness	-0,02	0,02	[-0,06; 0,02]
IAJ	-0,07	0,03	[-0,12; -0,01]
Einstiegsalter * IAJ	0,04	0,01	[0,01; 0,06]

Der Parameter ‚intercept‘ bezieht sich auf den mittleren G-Faktorenwert der Teilnehmenden zum ersten Erhebungszeitpunkt (t0) und der ‚slope‘-Parameter ist die Schätzung der mittleren Wachstumsrate im G-faktor der Teilnehmenden. Durch die ‚variances‘ werden die Varianz des G-Faktors zum Zeitpunkt t und die Residualvarianzen des ‚intercepts‘ und der ‚slope‘ geschätzt. Alle Prädiktoren die als Kovariaten mit den latenten Wachstumskoeffizienten modelliert wurden, sind stellen die Werte des ersten Erhebungszeitpunkts da und sind damit TAD-Konform.  $CFI = 0,98$ ,  $TLI = 0,95$ ,  $\chi^2(10,356) = 13,43$  mit  $p = 0,20$ ,  $RMSEA = 0,03$  mit  $95\%CI [0; 0,07]$ ,  $SRMR = 0,02$

**Förderung** Das Forschungsprojekt wird durch Mittel aus dem Anneliese-Maier-Forschungspreis der Alexander von Humboldt-Stiftung unterstützt.

**Funding** Open access funding provided by Goldsmiths College.

**Interessenkonflikt** P. Labonde und D. Müllensiefen erklären, dass keine konkurrierenden Interessen bestehen.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

## Literatur

- Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. H. Bower (Hrsg.), *The psychology of learning and motivation* (S. 47–90). London: Academic Press.
- Bastian, H. G., & Koch, M. (2010). *Vom Karrieretraum zur Traumkarriere – Eine Langzeitstudie über musikalische Hochbegabte*. Mainz: Schott.
- Bella, D. S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: battery for the assessment of auditory sensorimotor and timing abilities. *Behavior Research Methods*, 49(3), 1128–1145. <https://doi.org/10.3758/s13428-016-0773-6>.
- Bentley, A. (1966). *Measures of musical abilities*. London: Harrap.
- Berz, W. A. (1995). Working memory in music: A theoretical model. *Music Perception*, 12(3), 353–364. <https://doi.org/10.2307/40286188>.
- Biddle, S. J. H., Wang, C. K. J., Chatzisarantis, N. L. D., & Spray, C. M. (2003). Motivation for physical activity in young people: entity and incremental beliefs about athletic ability. *Journal of Sports Sciences*, 21(12), 973–989. <https://doi.org/10.1080/02640410310001641377>.
- Bullerjahn, C., Dziewas, J., Hilsdorf, M., Kassel, C., Menze, J., & Gembris, H. (2020). Why adolescents participate in a music contest and why they practice—the influence of incentives, flow, and volition on practice time. *Frontiers in Psychology*, 11, 2771. <https://doi.org/10.3389/fpsyg.2020.561814>.
- Bürkner, P. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>.
- Bürkner, P. (2021). Handle Missing Values with brms. CRAN. [https://cran.r-project.org/web/packages/brms/vignettes/brms\\_missings.html](https://cran.r-project.org/web/packages/brms/vignettes/brms_missings.html). Zugegriffen: 4. Juli 2022.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67.
- Chan, Y. W. F., & Kosinski, M. (2015). *ICAR Project Wiki*. International Cognitive Ability Resource (ICAR). <https://icar-project.com/projects/icar-project/wiki>. Zugegriffen: 4. Juli 2022.
- Clausen, B., & Lessing, W. (2018). Institutionen des Musiklernens im deutschsprachigen Raum. In M. Dartsch, J. Knigge, A. Nießen, F. Platz & C. Stöger (Hrsg.), *Handbuch Musikpädagogik: Grundlagen – Forschung – Diskurse* (S. 387–396). Münster: Waxmann.
- Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, 43, 52–64. <https://doi.org/10.1016/j.intell.2014.01.004>.
- Creech, A., Papageorgi, I., Duffy, C., Morton, F., Hadden, E., Potter, J., De Bezenac, C., Whyton, T., Himonides, E., & Welch, G. (2008). Investigating musical performance: commonality and diversity among classical and non-classical musicians. *Music Education Research*, 10(2), 215–234. <https://doi.org/10.1080/14613800802079080>.
- Deutsches Musikinformationszentrum (2021). *Jugendorchester, -ensembles und -chöre*. MIZ. <http://www.miz.org/institutionen/jugendorchester-jugendensembles-jugendchoere-s25>. Zugegriffen: 4. Juli 2022.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101.
- Dweck, C. S. (2000). *Self-theories: their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: a word from two perspectives. *Psychological Inquiry*, 6(4), 267–285. [https://doi.org/10.1207/s15327965pli0604\\_1](https://doi.org/10.1207/s15327965pli0604_1).
- Eisinger, M. (2021). Struktur Subjektiver Theorien über Musikalität und Validierung eines Messinstruments. *Jahrbuch Musikpsychologie*, 30, Artikel e93. <https://doi.org/10.5964/jbdgm.93>.

- van Erp, S. (2020). A tutorial on Bayesian penalized regression with shrinkage priors for small sample sizes. In R. van de Schoot & M. Miočević (Hrsg.), *Small sample size solutions: a guide for applied researchers and practitioners* (S. 71–84). New York: Routledge.
- Fiedler, D., & Müllensiefen, D. (2015). Validation of the Gold-MSI questionnaire to measure musical sophistication of German students at secondary education schools. *Research in Music Education*, 36, 199–219.
- Fujii, S., & Schlaug, G. (2013). The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception and production and their dissociation. *Frontiers in Human Neuroscience*, 7, 771. <https://doi.org/10.3389/fnhum.2013.00771>.
- Furnham, A., Jensen, T., & Crump, J. (2008). Personality, Intelligence and Assessment Centre Expert Ratings. *International Journal of Selection and Assessment*, 16, 356–365. <https://doi.org/10.1111/j.1468-2389.2008.00441.x>.
- Gagné, F. (1998). A proposal for subcategories within gifted or talented populations. *Gifted Child Quarterly*, 42(2), 87–95. <https://doi.org/10.1177/001698629804200203>.
- Gagné, F., & McPherson, G. E. (2016). Analyzing musical prodigiousness using Gagné’s integrative model of talent development. In G. E. McPherson (Hrsg.), *Musical prodigies* (S. 3–114). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685851.003.0001>.
- Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences*. London: Macmillan and Co.
- Geddes, B. (1990). How the cases you choose affect the answers you get: selection bias in comparative politics. *Political Analysis*, 2, 131–150. <https://doi.org/10.1093/pan/2.1.131>.
- Gembris, H. (2002). *Grundlagen musikalischer Begabung und Entwicklung* (2. Aufl.). Augsburg: Wißner.
- Ghisletta, P., Cantoni, E., & Jacot, N. (2015). Nonlinear growth curve models. In M. Stemmler, A. von Eye & W. Wiedermann (Hrsg.), *Dependent data in social sciences research* (S. 47–66). Springer. [https://doi.org/10.1007/978-3-319-20585-4\\_2](https://doi.org/10.1007/978-3-319-20585-4_2).
- Goldberg, L. R. (1990). An alternative “description of personality”: the Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216–1229. <https://doi.org/10.1037/0022-3514.59.6.1216>.
- Goldstein, H., & French, R. (2015). Differential educational progress and measurement error. *Longitudinal and Life Course Studies*, 6(3), 3351–3358.
- Gordon, E. E. (1989). *Advanced measures of music audition*. Chicago: GIA.
- Gordon, E. E. (1995). *Musical aptitude profile*. Chicago: GIA.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1).
- Harrison, P. M. C., & Müllensiefen, D. (2018). Development and validation of the computerised adaptive beat alignment test (CA-BAT). *Scientific Reports*, 8(1), 12395. <https://doi.org/10.1038/s41598-018-30318-8>.
- Harrison, P. M. C., Collins, T., & Müllensiefen, D. (2017). Applying modern psychometric techniques to melodic discrimination testing: Item response theory, computerised adaptive testing, and automatic item generation. *Scientific Reports*, 7(1), 3618. <https://doi.org/10.1038/s41598-017-03586-z>.
- Heller, K. A., & Perleth, C. (2005). *Münchener Hochbegabungsbatterie für die Primarstufe (MHB-P) und Sekundarstufe (MHBT-S)*. Göttingen: Hogrefe.
- Hespanhol, L., Vallio, C. S., Costa, L. M., & Saragiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Brazilian Journal of Physical Therapy*, 23(4), 290–301. <https://doi.org/10.1016/j.bjpt.2018.12.006>.
- Hochschule für Musik Würzburg (2017). Studien- und Prüfungsordnung des Pre-College Würzburg. [https://www.hfm-wuerzburg.de/amt/Studien-\\_Pruefungsordnung/PC/PreCollege\\_Wuerzburg\\_Studien\\_und\\_Pruefungsordnung\\_2017.05.24.pdf](https://www.hfm-wuerzburg.de/amt/Studien-_Pruefungsordnung/PC/PreCollege_Wuerzburg_Studien_und_Pruefungsordnung_2017.05.24.pdf). Zugegriffen: 4. Juli 2022.
- Hodges, D. A., & Haak, P. A. (1996). The influence of music on human behavior. In D. A. Hodges (Hrsg.), *Handbook of music psychology* (S. 469–555). San Antonio: IMR Press.
- Homan, M. D., & Gelman, A. (2014). The no-U-turn sampler: adaptively setting path lengths in hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1), 1593–1623.
- Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, 1252(1), 85–91. <https://doi.org/10.1111/j.1749-6632.2011.06402.x>.
- Iversen, J. R., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima & M. Tsuzaki (Hrsg.), *Proceedings of the 10th International Conference on Music Perception and Cognition* (S. 465–468). Sapporo: ICMPC.

- Jerrim, J., & Vignoles, A. (2013). Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(4), 887–906. <https://doi.org/10.1111/j.1467-985X.2012.01072.x>.
- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302–317. <https://doi.org/10.1111/j.1751-9004.2007.00054.x>.
- Junge Deutsche Philharmonie (2021). *Das Zukunftsorchester*. Jdph. <https://www.jdph.de/de/jdph/biografie>. Zugegriffen: 4. Juli 2022.
- Karma, K. (1984). Musical aptitude as the ability to structure acoustic material. *International Journal of Music Education*, 3(1), 27–30. <https://doi.org/10.1177/025576148400300104>.
- Larrouy-Maestri, P., Harrison, P. M. C., & Müllensiefen, D. (2019). The mistuning perception test: a new measurement instrument. *Behavior Research Methods*, 51(2), 663–675. <https://doi.org/10.3758/s13428-019-01225-1>.
- Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: construction and validation of the profile of music perception skills. *PLoS ONE*, 7(12), e52508. <https://doi.org/10.1371/journal.pone.0052508>.
- Lin, H., Turnbull, B. W., McCulloch, C. E., & Slate, E. H. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data. *Journal of the American Statistical Association*, 97(457), 53–65. <https://doi.org/10.1198/016214502753479220>.
- MacGregor, C., & Müllensiefen, D. (2019). The musical emotion discrimination task: a new measure for assessing the ability to discriminate emotions in music. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2019.01955>.
- Manturzewska, M. (1990). A biographical study of the life-span development of professional musicians. *Psychology of Music*, 18(2), 112–139. <https://doi.org/10.1177/0305735690182002>.
- Mayr, T., Bauer, C., & Krause, M. (2010). *KOMPIK – Kompetenzen und Interessen von Kindern. Beobachtungs- und Einschätzungsbögen für Kinder von 3,5 bis 6 Jahre*. Gütersloh: Bertelsmann.
- McIntosh, D. E., Dixon, F. A., & Pierson, E. E. (2018). Use of intelligence tests in the identification of giftedness. In D. P. Flanagan & E. M. McDonough (Hrsg.), *Contemporary intellectual assessment: Theories, tests, and issues* (S. 587–607). New York: Guilford.
- McPherson, G. E., & Williamon, A. (2016). Building gifts into musical talents. In G. E. McPherson (Hrsg.), *The child as musician: a handbook of musical development* (S. 340–360). Oxford: Oxford University Press.
- Michaelis, K. F. (1805). Über die Prüfung musikalischer Fähigkeiten. *Berlinische Musikalische Zeitung*, 56–58, 222–230.
- Müllensiefen, D. (2017). Messung musikalischer Begabung. In U. Trautwein & M. Hasselhorn (Hrsg.), *Begabungen und Talente* (S. 125–144). Göttingen: Hogrefe.
- Müllensiefen, D., & Hemming, J. (2018). Musikalische Fertigkeiten und ihre Messbarkeit. In A. C. Lehmann & R. Kopiez (Hrsg.), *Handbuch Musikpsychologie* (S. 93–119). Bern: Hogrefe.
- Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. (2014). The musicality of non-musicians: an index for measuring musical sophistication in the general population. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0089642>.
- Müllensiefen, D., Harrison, P., Caprini, F., & Fancourt, A. (2015). Investigating the importance of self-theories of intelligence and musicality for students' academic and musical achievement. *Frontiers in Psychology*, 6, 1702. <https://doi.org/10.3389/fpsyg.2015.01702>.
- Müllensiefen, D., Kozbelt, A., Olszewski-Kubilius, P., Subotnik, R., Worrell, F. C., & Preckel, F. (2022). Talent development in music. In G. McPherson (Hrsg.), *The Oxford Handbook of Music Performance*, Vol. 1. (S. 85–106). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190056285.013.32>
- Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J., & Schlaug, G. (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain and Cognition*, 59(2), 124–134. <https://doi.org/10.1016/j.bandc.2005.05.009>.
- Pape, W. (2013). Familiäre musikalische Sozialisation. In R. Heyer, S. Wachs & C. Palentien (Hrsg.), *Handbuch Jugend – Musik – Sozialisation* (S. 219–248). Berlin: Springer. [https://doi.org/10.1007/978-3-531-18912-3\\_6](https://doi.org/10.1007/978-3-531-18912-3_6).
- Pausch, V., Müllensiefen, D., & Kopiez, R. (2022). Musikalischer g-Faktor oder multiple Faktoren? Struktur und Leistungskennwerte der musikalischen Hörfähigkeit von Jugendlichen. *Jahrbuch Musikpsychologie*, 30, Artikel e89. <https://doi.org/10.5964/jbdgm.89>.

- Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*(2), 309–322. <https://doi.org/10.1037/0033-2909.84.2.309>.
- Preckel, F., Golle, J., Grabner, R., Jarvin, L., Kozbelt, A., Müllensiefen, D., Olszewski-Kubilius, P., Schneider, W., Subotnik, R., Vock, M., & Worrell, F. C. (2020). Talent development in achievement domains: a psychological framework for within- and cross-domain research. *Perspectives on Psychological Science*, *15*(3), 691–722. <https://doi.org/10.1177/1745691619895030>.
- Proust-Lima, C., Philipps, V., & Liqueur, B. (2017). Estimation of extended mixed models using latent classes and latent processes: the R package lamm. *Journal of Statistical Software*, *78*(2), 1–56.
- Rademacher, U. (2017). Fördern und fordern bei Jugend musiziert. *Beiträge Zur Hochschulpolitik*, *1*, 12–16.
- Revelle, W. (2021). *psych: Procedures for Psychological, Psychometric, and Personality Research* [Computer Programm] (Ver. 2.1.6). <https://cran.r-project.org/package=psych>. Zugegriffen: 4. Juli 2022.
- Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2018). The role of posterior parietal cortex in Beat-based timing perception: a continuous theta burst stimulation study. *Journal of Cognitive Neuroscience*, *30*(5), 634–643. [https://doi.org/10.1162/jocn\\_a\\_01237](https://doi.org/10.1162/jocn_a_01237).
- Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>.
- Ruth, N., & Müllensiefen, D. (2021). Survival of musical activities. When do young people stop making music? *PLoS ONE*, *16*(11), e259105. <https://doi.org/10.1371/journal.pone.0259105>.
- Sala, G., & Gobet, F. (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, *20*, 55–67. <https://doi.org/10.1016/j.edurev.2016.11.005>.
- Schaal, N. K., Bauer, A. K. R., & Müllensiefen, D. (2014). Der Gold-MSI: Replikation und Validierung eines Fragebogeninstrumentes zur Messung Musikalischer Erfahrung anhand einer deutschen Stichprobe. *Musicae Scientiae*, *18*(4), 423–447. <https://doi.org/10.1177/1029864914541851>.
- Seashore, C. E. (1919). *Seashore measurement of musical talent*. New York: Schirmer.
- Silas, S., Müllensiefen, D., Gelding, R., Frieler, K., & Harrison, P. (2022). The associations between music training, musical working memory, and visuospatial working memory: An opportunity for causal modelling. *Music Perception*, *39*(4), 401–420. <https://doi.org/10.1525/mp.2022.39.4.401>.
- Sloboda, J. A., & Howe, M. J. A. (1991). Biographical precursors of musical excellence: an interview study. *Psychology of Music*, *19*(1), 3–21. <https://doi.org/10.1177/0305735691191001>.
- Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: a proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, *12*(1), 3–54. <https://doi.org/10.1177/1529100611418056>.
- Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2012). A proposed direction forward for gifted education based on psychological science. *Gifted Child Quarterly*, *56*(4), 176–188. <https://doi.org/10.1177/0016986212456079>.
- Tsigeman, E., Silas, S., Frieler, K., Likhanov, M., Gelding, R., Kovas, Y., & Müllensiefen, D. (2022). The Jack and Jill adaptive working memory task: construction, calibration and validation. *PLoS ONE*, *17*(1), e262200. <https://doi.org/10.1371/journal.pone.0262200>.
- Ullén, F., Mosing, M., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, *63*, 87–93. <https://doi.org/10.1016/j.paid.2014.01.057>.
- Ullén, F., Hambrick, D. Z., & Mosing, M. A. (2016). Rethinking expertise: A multifactorial gene-environment interaction model of expert performance. *Psychological Bulletin*, *142*(4), 427–446. <https://doi.org/10.1037/bul0000033>.
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The musical ear test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*, 188–196. <https://doi.org/10.1016/J.LINDIF.2010.02.004>.
- Wing, H. D. (1939). *Standardised tests of musical intelligence*. The Mere England: National Foundation for Educational Research.
- Winner, E., & Martino, G. (2000). Giftedness in non-academic domains: the case of the visual arts and music. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Hrsg.), *International handbook of giftedness and talent* (S. 95–110). New York: Elsevier.
- Wood, S. N. (2017). *Generalized additive models: an introduction with R* (2. Aufl.). Boca Raton: Chapman and Hall.