

Wahrnehmungsräume melodischer Varianten bei Schüler:innen: Eine multidimensionale Analyse globaler Ähnlichkeitsurteile

Perceptual Spaces of Melodic Variation in School Students: A Multidimensional Analysis of Global Similarity Judgments

Myrto Zarzalis¹ , Johannes Hasselhorn¹ , Daniel Fiedler² 

[1] Lehrstuhl für Musikpädagogik und -didaktik, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Nürnberg, Deutschland. [2] Institut für Musikwissenschaft, Musikpädagogik und Ästhetik, Staatliche Hochschule für Musik und Darstellende Kunst Stuttgart (HMDK), Stuttgart, Deutschland.

Jahrbuch Musikpsychologie, 2026, Vol. 34, Article e255, <https://doi.org/10.5964/jbdgm.255>

Received: 2026-06-15 • **Accepted:** 2026-10-16 • **Published (VoR):** 2026-10-29

Reviewed by: Klaus Frieler; Felix Thiesen.

Corresponding Author: Myrto Zarzalis, Lehrstuhl für Musikpädagogik und -didaktik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Regensburger Straße 160, 90478 Nürnberg, Deutschland. E-mail: myrto.zarzalis@fau.de

Supplementary Materials: Data, Materials [see [Index of Supplementary Materials](#)]



Zusammenfassung

Die auditive Wahrnehmung musikalischer Strukturen bildet eine wesentliche Grundlage des musikalischen Verstehens. Obwohl unterschiedliche melodische Variantenarten (wie beispielsweise Sequenzierung, Dynamik oder rhythmische Varianten) musiktheoretisch kategorisiert werden können, ist bislang unzureichend geklärt, wie Menschen solche Variantenarten aufgrund ihrer wahrgenommenen Ähnlichkeit organisieren und inwiefern diese Wahrnehmungsstruktur musiktheoretischen Kategorien entspricht. Die vorliegende Studie untersuchte, ob sich aus den Ähnlichkeitsurteilen von Schüler:innen ein konsistenter Wahrnehmungsraum melodischer Variantenarten rekonstruieren lässt und welche Strukturen sich darin zeigen. Hierzu wurden Ähnlichkeitsentscheidungen von 1.336 Schüler:innen der Klassenstufen 5 bis 11 zu systematisch variierten Tonfolgen erhoben und mittels Multidimensionaler Skalierung und Clusteranalyse ausgewertet. Die Ergebnisse deuteten darauf hin, dass die wahrnehmungsseitige Organisation melodischer Ähnlichkeit in ihrer Grundstruktur Bezüge zu musiktheoretischen Differenzierungen aufwies und die Variantenarten zugleich in einem mehrdimensionalen System organisiert waren. Die erste Dimension war durch eine Trennung zeitlich-rhythmischer und tonhöhenbasierter



This is an open access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Variantenarten gekennzeichnet, während die zweite Dimension eine Gegenüberstellung performativ-klanglicher und relational-struktureller Variantenarten erkennen ließ. Die dritte Dimension wies hingegen keine eindeutige inhaltliche Zuordnung auf. Darüber hinaus deuteten die Ergebnisse auf mögliche Expertiseabhängigkeiten bei der Strukturierung der Ähnlichkeitsurteile hin. Insgesamt bietet die Studie eine Grundlage dafür, musiktheoretische Konzepte wie musikalische Parameter im schulischen Musikunterricht stärker an wahrnehmungsbasierte Organisationsstrukturen der Schüler:innen anzubinden.

Schlüsselwörter

Auditive Strukturwahrnehmung, Mental-auditive Repräsentationen, Musikalische Parameter, Musikalische Expertise, Clusteranalyse, Musikpädagogik

Abstract

Auditory perception of musical structures forms an essential basis for musical understanding. Although music theory can categorize different types of melodic variation (such as sequential, dynamic, or rhythmic variation), it remains insufficiently understood how individuals organize these variations based on perceived similarity and to what extent this perceptual structure corresponds to music-theoretical categories. The present study investigated whether a consistent perceptual space of melodic variation types can be reconstructed from students' similarity judgments and which structures emerge within this space. To this end, similarity judgments from 1,336 students in grade levels 5 to 11 were collected for systematically varied tone sequences and analyzed using multidimensional scaling and cluster analysis. The results suggested that the perceptual organization of melodic similarity was broadly related to music-theoretical distinctions, while the types of variation were simultaneously organized within a multidimensional system. The first dimension distinguished between temporal-rhythmic and pitch-based variations; the second contrasted performance-related/timbral and relational-structural variations; and the third did not permit a clear substantive interpretation. In addition, the results suggested potential expertise-related differences in the organization of similarity judgments. Overall, the study provides a basis for more closely aligning music-theoretical concepts, such as musical parameters, with students' perceptual organization in school music education.

Keywords

auditory structure perception, mental auditory representations, musical parameters, musical expertise, cluster analysis, music education

Die Erkennung und die Verarbeitung musikalischer Strukturen sind grundlegende menschliche kognitive Fähigkeiten und Fertigkeiten, die sowohl das musikalische Verstehen als auch die ästhetische Erfahrung beeinflussen können (Altenmüller, 2018). Musikalische Strukturen beruhen auf der Organisation von Elementen und ihren Relationen, wobei insbesondere bei Schüler:innen interindividuelle Unterschiede in der Wahrnehmung dieser Relationen anzunehmen sind. Solche interindividuellen Unterschiede

könnten dabei zu den von Meyer (2024) beobachteten teils stark divergierenden Beschreibungen derselben Musik durch Schüler:innen beitragen.

Somit stellt sich gerade für die Gestaltung von Lehr- und Lernumgebungen die übergeordnete Frage, inwiefern die individuelle Wahrnehmung von Schüler:innen mit musiktheoretisch definierten Strukturen übereinstimmt. Während die Musiktheorie musikalische Relationen kategorial systematisiert und analytisch beschreibt, untersucht die Musikpsychologie die perzeptiven und kognitiven Prozesse, auf deren Grundlage musikalische Ereignisse als ähnlich oder verschieden wahrgenommen werden können (Frierler, 2018). Bislang ist jedoch nur unzureichend empirisch geklärt, wie Schüler:innen melodische Variantenarten wahrnehmungsbasiert organisieren, inwiefern sich dabei musiktheoretische Klassifikationen widerspiegeln und inwieweit sich diese Organisation über unterschiedliche Gruppen hinweg als vergleichbar erweist. Vor diesem Hintergrund untersuchte die vorliegende Studie mittels Multidimensionaler Skalierung die Dimensionen der Ähnlichkeitswahrnehmung melodischer Variantenarten und deren Vergleichbarkeit zwischen Gruppen von Schüler:innen mit unterschiedlicher musikalischer Expertise. Mittels einer Clusteranalyse wurde zudem untersucht, inwiefern sich die musiktheoretisch klassifizierten Variantenarten in der Wahrnehmung widerspiegeln.

Stand der Forschung

Die Analyse musikalischer Strukturen bildet die Grundlage sowohl musiktheoretischer Diskurse als auch musikpädagogischer Zielsetzungen. Struktur wird im Allgemeinen als ein aus Elementen nach Regeln gebautes Ordnungsgefüge verstanden (Struktur (allgemein), o. D.). Im musikalischen Kontext können diese Elemente beispielsweise Tonfolgen oder Pattern entsprechen, während das Ordnungsgefüge als Aspekt des Formbegriffs (etwa Lied, Periode oder Rondo) bezeichnet werden kann (Kühn, 2021). Konstitutiv für das in der Definition beschriebene Ordnungsgefüge ist das Prinzip der Wiederholung, durch das sich Motive manifestieren und größere Formzusammenhänge entstehen (Götte, 2020; Kühn, 2021; Tenzer, 2006). Für die menschliche Wahrnehmung ist dabei nicht nur die exakte Wiederholung entscheidend, sondern insbesondere das Erkennen partieller Gleichheit, also das Erkennen von Varianten (Ahlbäck, 2007; Deliège, 2001).

Musiktheoretisch wird die Variante als Abwandlung definiert, bei der einzelne Merkmale konstant bleiben, während andere verändert werden (Kühn, 2021). Daraus resultieren klassische Kategorisierungen, beispielsweise rhythmische, tonhöhenbasierte oder harmonische Variantenarten (Dress & von Fischer, 2016; Schönberg, 1979). Jedoch stellt sich die Frage, inwieweit diese musiktheoretisch begründeten Klassifikationen der Variantenarten auch mit der subjektiven Hörerfahrung korrespondieren. Ähnlichkeit ist keine inhärente, objektive Eigenschaft akustischer Signale, sondern emergiert erst durch mental-auditive Repräsentationen und kognitive Organisationsprozesse (Hahn, 2014; Harrison et al., 2016). Unter mental-auditiven Repräsentationen werden interne, neurokognitive Abbilder externer Schallereignisse verstanden, durch die beispielsweise

die in der Definition musikalischer Strukturen beschriebenen musikalischen Elemente und ihre Relationen mental dargestellt werden können.

Zentral für die Entstehung dieser Repräsentationen ist das Wahrnehmungsencoding im auditorischen Kortex. Neurokognitive Studien zeigten dabei unterschiedliche Aktivierungsmuster bei tonhöhen- und zeitbasierten Inhalten (Lappe et al., 2013, 2016; Prince, 2014). Dementsprechend wird von einer Differenzierung ausgegangen, wonach tonhöhenbasierte und zeitliche Dimensionen unabhängig voneinander kognitiv verarbeitet werden.

Darüber hinaus unterliegt die Bildung mental-auditiver Repräsentationen einer Verarbeitungshierarchie, in der die Analyse elementarer Merkmale (Intensität, Tonhöhe) vor der Analyse komplexer Strukturen (wie Melodie, Harmonie oder Rhythmus) steht (Brattico et al., 2018; Jäncke, 2024; Koelsch & Schröger, 2018). Diese Hierarchie bietet eine plausible Erklärung dafür, dass in Ähnlichkeitsurteilen performativ-klangliche Oberflächenmerkmale wie Dynamik oder Artikulation vorrangig als Entscheidungsmerkmale herangezogen werden, wohingegen abstrakte Strukturparameter wie etwa Motivik oder Harmonik seltener berücksichtigt werden (Lamont & Dibben, 2001).

Zur Modellierung von Ähnlichkeitsurteilen lassen sich primär räumliche und merkmalsbasierte Ansätze unterscheiden. Räumliche Ansätze, die Ähnlichkeit als Distanz in einem psychologischen Raum konzeptualisieren (Shepard, 1962), wurden dahingehend kritisiert, dass ihre metrischen Annahmen kontextuelle Einflüsse, deren Bedeutung in zahlreichen Studien hervorgehoben wurde (Ahlbäck, 2007; Cambouropoulos, 2009; Taher et al., 2018), und asymmetrische Urteile häufig unzureichend berücksichtigen (Hahn, 2014). Dies führte zur Entwicklung merkmalsbasierter Ansätze wie dem Kontrastmodell (Tversky, 1977), das Ähnlichkeit als Funktion gewichteter gemeinsamer und unterscheidender Merkmale beschreibt. Die Relevanz räumlicher Ansätze wurde durch diese Kritik jedoch nicht grundsätzlich infrage gestellt: Krumhansl (1990) konnte zeigen, dass unter kontrollierten Bedingungen, etwa bei gleicher Tonalität und vergleichbarer Stimulilänge, die Ähnlichkeitsurteile annähernd symmetrisch sind und sich somit robust als Distanzen abbilden lassen. Vor diesem Hintergrund etablierte sich die Multidimensionale Skalierung (MDS) als eine geeignete Methode zur empirischen Rekonstruktion solcher Wahrnehmungsräume und ermöglichte die Identifikation zentraler Dimensionen wie Kontur und Ambitus (Eerola & Bregman, 2007) sowie Tempo, Genre und Timbre (Novello et al., 2011) bei der Modellierung von Ähnlichkeitswahrnehmung.

Da Ähnlichkeitsurteile maßgeblich von mental-auditiven Repräsentationen abhängen, kann musikalische Expertise (Hasselhorn & Knigge, 2026) den Urteilsprozess beeinflussen. Expertisierung erfolgt dabei in der Regel erfahrungsbasiert. In Bezug auf Leistungen im Bereich des praktischen instrumentalen bzw. vokalen Musizierens stellt eine kumulierte Übendauer eine gute Annäherung an dafür relevante Lebenserfahrung dar (Gruber & Lehmann, 2014; Platz et al., 2014). Für Stichproben im Altersbereich von Schüler:innen werden daher häufig die Teilnahme an außerschulischem Musikunter-

richt und das Alter als Indikatoren für akkumulierte musikbezogene Hör- und Lernerfahrungen herangezogen, um Unterschiede in der musikalischen Expertisierung bei der Wahrnehmung melodischer Varianten näherungsweise zu operationalisieren (Hasselhorn & Knigge, 2026). Dabei kann angenommen werden, dass sich ein möglicher Einfluss von außerschulischem Vokal- und Instrumentalunterricht in der Struktur von Ähnlichkeitsurteilen zeigt: Während der Urteilsprozess bei weniger expertisierten Hörer:innen tendenziell von Oberflächenmerkmalen wie Kontur, Tempo oder Timbre dominiert wird (Lamont & Dibben, 2001), sind stärker expertisierte Hörer:innen zunehmend in der Lage, auch abstrakte strukturelle Invarianten als primäre Bewertungsgrundlage heranzuziehen (Eitan & Granot, 2009).

Ein möglicher Erklärungsansatz für diese Diskrepanz besteht darin, dass musikalische Expertisierung zu einer differenzierteren Wahrnehmung regelbasierter Abweichungen beiträgt. Zwar verfügen alle Menschen aufgrund von Mechanismen des statistischen Lernens über ein implizites Wissen hinsichtlich musikalischer Regularitäten (Rohrmeier & Rebuschat, 2012), doch ist dieses bei musikalischen Expert:innen deutlich differenzierter ausgebildet, was sich in präziseren Erwartungshaltungen und einer effizienteren Informationsverarbeitung zeigt (Bonetti et al., 2022; Zhao et al., 2017). Dass Expert:innen strukturelle Invarianten im Ähnlichkeitsvergleich stärker gewichten, korrespondiert direkt mit ihren neuronal verankerten Fähigkeiten und Fertigkeiten, Abweichungen von solchen strukturellen Regelmäßigkeiten auch in komplexen musikalischen Zusammenhängen zu erkennen und für die Urteilsbildung nutzbar zu machen (Bonetti et al., 2022; van Zuijen et al., 2004, 2005).

Im schulischen Kontext spiegelt die curriculare Progression von der Identifikation isolierter Parameter (z.B. Klassenstufen 5/6) hin zur Analyse relationaler Strukturen (ab Klassenstufen 7/8; siehe hierzu [Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016](#)) diese Entwicklung und korrespondiert mit der entwicklungsbedingten Zunahme der Arbeitsgedächtniskapazität und der Abstraktionsfähigkeit von Schüler:innen (Cowan, 2016; Jentschke & Koelsch, 2006). Dabei kann die Verarbeitung komplexerer musikalischer Zusammenhänge auf basalen perzeptiven Fähigkeiten wie der Wahrnehmung von Tonhöhen und Tonhöhenveränderungen aufbauen, die bereits vor Beginn der Sekundarstufe ein erwachsenenähnliches Niveau erreichen (Degé & Roden, 2018; Thompson, 2009). Inwiefern diese curriculare Progression auch mit einer entsprechenden Entwicklung der auditiven Strukturwahrnehmung einhergeht, ist jedoch offen. Empirisch-qualitative Befunde deuten darauf hin, dass Musiklehrkräfte musiktheoretische Inhalte oftmals rein terminologisch und ohne auditive Verankerung vermitteln (Ruf, 2024). Damit fehlt häufig eine unmittelbare Verknüpfung zwischen musiktheoretischem Begriff und auditiver Wahrnehmung.

Ziel der Studie

Für den schulischen Musikunterricht, in dem die Förderung des auditiven Strukturerkennens als zentrales Bildungsziel des Fachs Musik verankert ist (Altenmüller, 2018; Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016), stellt sich die Frage, nach welchen Kriterien Schüler:innen melodische Varianten perceptiv organisieren. Auf Grundlage des aktuellen Forschungsstandes wird angenommen, dass Prozesse der Ähnlichkeitswahrnehmung eine zentrale Voraussetzung für das Erkennen musikalischer Struktur bilden. Wie dargelegt, resultiert musikalische Ähnlichkeitswahrnehmung nicht ausschließlich aus objektiven Eigenschaften akustischer Ereignisse, sondern auch aus deren Verarbeitung vor dem Hintergrund individueller mentaler Repräsentationen und kognitiver Organisationsprozesse (Hahn, 2014; Harrison et al., 2016). Zwar existieren bereits zahlreiche Studien zur Wahrnehmung von Ähnlichkeit in den Bereichen Melodie (Eerola & Bregman, 2007), Tonalität (Krumhansl, 1979) und Rhythmus (Cao et al., 2014) sowie zur Ähnlichkeitswahrnehmung von Melodien unter Einbezug verschiedener musikalischer Merkmale (Müllensiefen & Frieler, 2007). Jedoch wurde bislang noch nicht empirisch untersucht, wie jene Variantenarten im Vergleich wahrnehmungsseitig organisiert werden, die explizit Gegenstand der schulischen Curricula sind. Es bleibt folglich offen, in welchem Verhältnis die musiktheoretischen Klassifikationen dieser Variantenarten zu den tatsächlichen perceptiven Organisationsdimensionen stehen, entlang derer Schüler:innen Ähnlichkeit beurteilen.

Vor diesem Hintergrund untersuchte die vorliegende Studie die Wahrnehmung melodischer Variantenarten bei Schüler:innen anhand von Ähnlichkeitsurteilen zu systematisch erzeugten Varianten von Tonfolgen. Das Ziel war es, einen Wahrnehmungsraum melodischer Variantenarten zu identifizieren, dessen Struktur hinsichtlich ihrer Vergleichbarkeit zwischen Gruppen unterschiedlicher musikalischer Expertise zu prüfen und die resultierenden Gruppierungen hinsichtlich ihrer Übereinstimmung mit musiktheoretischen Kategorien zu untersuchen. Daraus ergaben sich folgende Forschungsfragen (FF):

- FF1: Ergeben sich aus der Wahrnehmung von Variantenarten melodischer Tonfolgen von Schüler:innen – auch unter Berücksichtigung musikalischer Expertise – konsistente Dimensionen eines Wahrnehmungsraums?
- FF2: Entsprechen die Gruppierungen (Cluster) der Variantenarten melodischer Tonfolgen im Wahrnehmungsraum musiktheoretischen Klassifikationen?

Methode

Stichprobe

Die Stichprobe umfasste 1.336 Schüler:innen der Klassenstufen 5 bis 11 an Gymnasien in Baden-Württemberg, die im Zeitraum von April bis Juni 2025 an der Studie teilnahmen, wobei 832 Schüler:innen der Unterstufe (Klassenstufen 5–7) und 504 der Mittel- bzw. Oberstufe (Klassen 8–11) angehörten (siehe [Tabelle 1](#)). Hinsichtlich des Geschlechts identifizierten sich 606 Schüler:innen als weiblich, 712 als männlich und 18 als divers.

Tabelle 1

Verteilung der Stichprobe nach Jahrgangsstufen und außerschulischem Instrumental- oder Vokalunterricht

Klassenstufe	Außerschulischer Instrumental- oder Vokalunterricht	
	ja	nein
Unterstufe		
5	155	136
6	140	122
7	128	151
Mittelstufe		
8	100	126
9	40	86
10	64	79
11	6	3
Gesamt	633	703

Anmerkung. Aufgrund der geringen Anzahl an Schüler:innen aus der Jahrgangsstufe 11 wurden diese der Mittelstufe zugeordnet.

Ablauf der Erhebung

Die Studie fand im schulischen Musikunterricht während der regulären Unterrichtszeit statt. Die Datenerhebung erfolgte mithilfe der Online-Plattform SoSci Survey (Version 3.5.02; [Leiner, 2024](#)). Der eingesetzte Fragebogen umfasste demografische Angaben sowie 15 selbst erstellte Aufgaben zur Erfassung von Ähnlichkeitsentscheidungen, die durch zehn optionale Bonusaufgaben ergänzt wurden. Die Schüler:innen bearbeiteten die Aufgaben einzeln an Tablets oder PCs und nutzten dabei Kopfhörer. Es wurde jeweils die an der Schule vorhandene technische Ausstattung verwendet, deren Funktionsfähigkeit vor Ort von der betreuenden Lehrkraft überprüft wurde. Die gesamte Erhebung, einschließlich des technischen Aufbaus, der Instruktion und der Durchführung, dauerte pro Klasse etwa 30 Minuten.

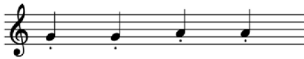
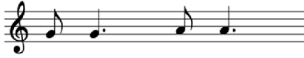
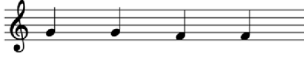
Stimulusmaterial

Als Stimulusmaterial dienten fünf viertönige Tonfolgen, die zufällig aus einer bestehenden Korpusstudie (Zarzalís et al., 2026) zu Tonfolgen im schulischen Musikunterricht ausgewählt wurden (Positionen 4, 7, 17, 28 und 39 im Korpus; Supplement A; <https://osf.io/cfa83>). Die Begrenzung auf Viertonfolgen erfolgte unter Berücksichtigung der Anforderungen an das Arbeitsgedächtnis (Cowan, 2001; Frieler, 2018; Miller, 1956) sowie auf Grundlage vorbereitender Erprobungen der im Rahmen dieser Studie verwendeten Triple-Vergleichsaufgaben. Für alle Tonfolgen galt ein einheitliches Grundtempo von Viertelnoten = 170 bpm.

Die Tonfolgen wurden mithilfe der Notationssoftware Capella (Version 10.0.6.0, Jungmann et al., 2024) als MIDI-Dateien realisiert und mit dem Klang „Konzertflügel“ wiedergegeben. Anschließend wurden die Stimuli als MP3-Dateien exportiert. Für jede der fünf ausgewählten Tonfolgen wurden zehn systematische Varianten (Abbildung 1) erstellt, die sich jeweils in einem musikalischen Parameter von der Originaltonfolge unterschieden. Die Varianten betrafen artikulatorische (Staccato), dynamische (Fortissimo, Crescendo), rhythmisch-zeitliche (Punktierung, irregulärer Rhythmus, Diminution) sowie tonhöhenbasierte Merkmale (Sequenz ↑, Sequenz ↓, Umkehrung, Intervalländerung). Art und Komplexität der Varianten orientierten sich an den im Bildungsplan Baden-Württemberg ausgewiesenen musikalischen Inhalten (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016). Diese musiktheoretisch definierten Merkmale dienten ausschließlich der systematischen Konstruktion des Stimulusmaterials und stellten keine vorgegebenen Wahrnehmungs- oder Bewertungsdimensionen für die Schüler:innen dar.

Abbildung 1

Übersicht über die zehn Varianten am Beispiel der Tonfolge 121 (Position 7 im Korpus)

Originale Tonfolge	
Fortissimo	
Crescendo	
Staccato	
Punktierung	
Irregulärer Rhythmus	
Diminution	
Sequenz ↑	
Sequenz ↓	
Umkehrung	
Intervalländerung	

Aufgabenstruktur

Die Beurteilungsaufgabe bestand aus einer Ähnlichkeitsentscheidung im Triple-Format (siehe [Abbildung 2](#) und Supplement B; <https://osf.io/cfa83>). Den Schüler:innen wurden jeweils drei Varianten derselben Tonfolge präsentiert, aus denen sie per Klick die beiden Hörbeispiele auswählen sollten, die sie als am ähnlichsten wahrnahmen. Die Wiedergabe

der Hörbeispiele innerhalb eines Triples war fremdgesteuert und erfolgte stets in derselben Präsentationsreihenfolge. Die Pausen zwischen den einzelnen Hörbeispielen wurden auf 1,5 Sekunden standardisiert, um vergleichbare Hörbedingungen zu gewährleisten.

Abbildung 2

Aufgabenstellung für die Ähnlichkeitsbeurteilung der Triple

Ähnlichkeit in Musik bewerten

Aufgabe 1 von 15

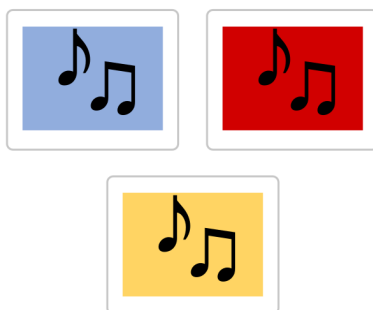
Im Folgenden hörst du jeweils drei kurze Melodien.

Klicke bitte die beiden Melodien an, die sich einander **am ähnlichsten** sind.*

Tippe auf den Lautsprecher, um die Hörbeispiele **anzuhören**.



Du kannst die Beispiele insgesamt dreimal hören.



* Wenn du dich bei der Eingabe vertippt haben solltest, klicke das Beispiel nochmals an, damit die Eingabe zurückgenommen wird.

Anmerkung. Für weitere Auszüge aus dem Fragebogen siehe Supplement B.

In der zentralen Beurteilungsphase bewerteten alle Schüler:innen insgesamt 15 Triples. Jede der fünf Tonfolgen erschien dabei dreimal, wobei die Zuordnung zu den Positionen 1 bis 15 randomisiert erfolgte. Nach Bearbeitung der 15 regulären Triples konnten die Schüler:innen freiwillig bis zu zwei weitere Blöcke mit jeweils fünf Triples bearbeiten. 727 Schüler:innen bearbeiteten mindestens den ersten, 380 auch den zweiten Bonusblock. Die Zuweisung der insgesamt 120 Triples pro Tonfolge erfolgte mithilfe der Urnen-Funktion von SoSci Survey (Version 3.5.02; [Leiner, 2024](#)), wobei für jede Tonfolge eine eigene Urne definiert wurde. Die Ziehung der Triples erfolgte ohne Zurücklegen über alle Schüler:innen hinweg, sodass eine möglichst gleichmäßige Verteilung der Triple-Konstellationen im Datensatz gewährleistet werden konnte. Sowohl die Auswahl der Tonfolge

und des Triples als auch die Abfolge der Varianten innerhalb eines Triples wurden zufällig bestimmt, um Reihen- und Positionseffekte zu minimieren.

Datenanalyse

Datenaufbereitung und Bildung der Ähnlichkeitsmaße

Die Datenanalysen basierten auf Ähnlichkeitsurteilen zu musikalischen Varianten, die auf der Auswahl der beiden ähnlichsten Elemente aus jeweils vorgelegten Triples beruhten. In jedem Triple wurden drei Varianten einer Tonfolge präsentiert, aus denen die Schüler:innen die beiden einander ähnlichsten Varianten auswählten. Eine einzelne Entscheidung entsprach damit der Auswahl eines von drei möglichen Variantenpaaren. In die weitere Auswertung ging jeweils das gewählte Variantenpaar ein. Die nicht gewählten Paare wurden nicht berücksichtigt.

Da die Bonusaufgaben freiwillig bearbeitet wurden, wurde ihr möglicher Einfluss mittels Sensitivitätsanalyse geprüft. Dabei zeigten sich bei Ausschluss der Bonusaufgaben keine substanziellen Veränderungen der MDS-Konfiguration (Prokrustes- $r = ,97$; PROTEST- $p < ,001$), sodass die bearbeiteten Bonusaufgaben in die weiteren Analysen einbezogen wurden. Über die gesamte Stichprobe hinweg wurde jedes Triple im Mittel 41,84-mal ($SD = 0,80$) bewertet. Die Verteilung der bearbeiteten Triples auf die Subgruppen kann [Tabelle 2](#) entnommen werden.

Tabelle 2

Mittlere Anzahl der Bewertungen pro Triple in der Gesamtstichprobe und den Subgruppen

Stichprobe	M	SD
Gesamte Stichprobe	41,84	0,80
Unterstufe	26,26	2,54
Mittelstufe	15,58	2,46
Instrumentalist:innen	19,87	3,23
Nichtinstrumentalist:innen	21,97	3,21

Zur Bildung der Proximitäten wurden die Ähnlichkeitswahlen für jedes Variantenpaar aggregiert. Da die kurzen, isolierten Tonfolgen unter kontrollierten Bedingungen und in randomisierter Reihenfolge innerhalb der Triples präsentiert wurden, wurde – gestützt durch Befunde annähernd symmetrischer Ähnlichkeitsurteile unter vergleichbaren Bedingungen (Krumhansl, 1990) – von einer annähernden Symmetrie der Urteile ausgegangen. Die Position der Varianten innerhalb eines Triples wurde deshalb bei der Aggregation nicht gesondert berücksichtigt. Für die weiteren Analysen wurden symmetrische Proximitätsmatrizen gebildet. Um Unterschiede in der Darbietungshäufigkeit der Paare, insbesondere zwischen den Subgruppen, zu berücksichtigen, wurden die Wahlhäufigkei-

ten normalisiert. Hierzu wurde die Anzahl der tatsächlichen Auswahl eines Paares durch die Anzahl der Situationen geteilt, in denen dieses Paar hätte gewählt werden können. Die resultierenden relativen Ähnlichkeitsmaße bildeten die Grundlage der nachfolgenden statistischen Analysen, die alle mit R (Version 4.5.2; [R Core Team, 2025](#)) durchgeführt wurden.

Prüfung der Vergleichbarkeit über die verschiedenen Tonfolgen

Um einen potenziellen Einfluss der fünf verschiedenen, eingesetzten Tonfolgen auf die Ähnlichkeitsbewertung ausschließen zu können, wurde zunächst deren tonfolgenübergreifende Vergleichbarkeit als Grundlage für die gemeinsame Modellierung geprüft. Hierzu wurden für jede Tonfolge zunächst separate ordinale Multidimensionale Skalierungen (MDS) auf Basis der vollständigen Stichprobe mit dem R-Paket *smacof* (Version 2.1.7; [Borg et al., 2018](#); [Mair et al., 2022](#)) berechnet. Die Güte der tonfolgen-spezifischen MDS-Lösungen wurde anhand der Stresswerte nach [Kruskal \(1964\)](#) beurteilt. Werte unter ,05 gelten in der Regel als exzellent, Werte unter ,10 als gut, wobei niedrigere Werte eine bessere Repräsentation der Disparitäten durch die Konfiguration anzeigen ([Borg & Groenen, 2005](#); [Kruskal, 1964](#)).

Die Vergleichbarkeit der tonfolgen-spezifischen Lösungen wurde anschließend auf zwei Ebenen untersucht. Die Übereinstimmung der Distanzrangordnungen der tonfolgen-spezifischen Lösungen wurde anhand von Spearman-Rangkorrelationen (ρ) der Disparitätsmatrizen erfasst. Werte über ,50 weisen auf einen starken Zusammenhang hin ([Cohen, 1988](#)), wobei höhere positive Koeffizienten auf ähnlichere Distanzrangordnungen hinweisen.

Die Übereinstimmung der räumlichen Konfigurationen wurde mittels Prokrustes-Analyse mit dem R-Paket *vegan* (Version 2.7-2; [Oksanen et al., 2025](#)) untersucht. Dabei wurden die Konfigurationen durch Rotation, Skalierung und Spiegelung optimal aneinander ausgerichtet. Die Prokrustes-Korrelation (r) diente als Maß der globalen Übereinstimmung. Zur Interpretation wurde das Quadrat dieses Wertes (r^2) als Anteil der erklärten Varianz der räumlichen Anordnung betrachtet ([Hair et al., 2019](#); [Legendre, 2012](#)). Die statistische Absicherung erfolgte mittels permutationsbasierter Prokrustes-Tests (PROTEST; $n = 999$; [Legendre, 2012](#)). Ergänzend wurde der Root Mean Square Error (RMSE) zur Beschreibung verbleibender Abweichungen der Punktpositionen herangezogen, um die lokale Präzision ergänzend zur globalen Strukturgüte einzuschätzen.

Zur ergänzenden inferenzstatistischen Prüfung tonfolgen-spezifischer Unterschiede in den Ähnlichkeitsentscheidungen wurden konditionale logistische Regressionsmodelle ([McFadden, 1974](#); [Train, 2009](#)) mit dem R-Paket *survival* (Version 3.8-9; [Therneau, 2026](#); [Therneau & Grambsch, 2000](#)) geschätzt. Jede Triple-Entscheidung wurde als Auswahl-situation mit drei möglichen Variantenpaaren modelliert, wobei das gewählte Paar mit 1 und die beiden nicht gewählten Paare mit 0 kodiert wurden. Die Abhängigkeit der drei Paaralternativen wurde durch die Stratifizierung nach Auswahl-situation berücksichtigt.

Wiederholte Entscheidungen derselben Person wurden durch clusterrobuste Standardfehler auf Personenebene berücksichtigt. Ein erstes Modell berücksichtigte Unterschiede in den Wahlwahrscheinlichkeiten zwischen den Variantenpaaren, während ein zweites Modell zusätzlich berücksichtigte, inwiefern diese Unterschiede von der zugrunde liegenden Tonfolge abhingen (Variantenpaar $\times$ Tonfolge). Diese Interaktion wurde mittels eines globalen clusterrobusten Wald-Tests geprüft. Die Modellgüte wurde anhand von McFaddens ρ^2 gegenüber einem Equal-Likelihood-Nullmodell mit einer Wahlwahrscheinlichkeit von 1/3 je Paar sowie der Veränderung zwischen beiden Modellen ($\Delta\rho^2$) beurteilt (McFadden, 1974).

Statistische Analysen

Zur Beantwortung von FF1 wurde auf Grundlage der aggregierten Ähnlichkeitsmaße mittels ordinaler Multidimensionaler Skalierung (MDS) untersucht, welche Dimensionen der Ähnlichkeitswahrnehmung der Variantenarten zugrunde lagen und inwiefern sich diese über Gruppen unterschiedlicher musikalischer Expertise hinweg vergleichbar darstellten. Die musiktheoretisch klassifizierten Variantenarten wurden dabei nicht als vorgegebene Dimensionen in die Modellierung einbezogen. Der Wahrnehmungsraum wurde ausschließlich auf Grundlage empirischer Ähnlichkeitsurteile rekonstruiert. Zur Bestimmung der Dimensionalität wurden für die Gesamtstichprobe ordinale MDS-Lösungen in Form gewichteter Individual Differences Scaling (INDSCAL-)Modelle mit zwei, drei und vier Dimensionen mit dem R-Paket *smacof* (Version 2.1.7; Mair et al., 2022) berechnet. Die Auswahl der Dimensionalität erfolgte anhand der Stresswerte sowie der relativen Stressreduktion zwischen den Lösungen. Die dimensionsspezifische Gewichtung berücksichtigte, dass Ähnlichkeitsurteile tonfolgeabhängig variieren können (Borg & Groenen, 2005; Carroll & Chang, 1970). Zur deskriptiven Charakterisierung des resultierenden Wahrnehmungsraums wurden die Streuung der Dimensionskoordinaten der Varianten sowie die normierten INDSCAL-Gewichte der Tonfolgen herangezogen.

Zur Untersuchung der Stabilität der Wahrnehmungsstruktur und ihrer Vergleichbarkeit zwischen den Subgruppen (siehe FF1) wurden separate MDS-Lösungen für die Subgruppen nach Schulstufe (Unter- vs. Mittel-/Oberstufe) sowie nach Instrumentalstatus (Schüler:innen mit vs. ohne außerschulischen Vokal- oder Instrumentalunterricht) berechnet. Die Modellpassung wurde zunächst anhand der Stresswerte beurteilt. Die globale Übereinstimmung der resultierenden Wahrnehmungsräume wurde nach Prokrustes-Ausrichtung anhand der Prokrustes-Korrelation beurteilt und mittels PROTEST ($n = 999$) geprüft (Legendre, 2012). Ergänzend wurde der RMSE herangezogen, um Abweichungen der Punktpositionen zu quantifizieren (Hair et al., 2019; Legendre, 2012). Zur näheren Lokalisierung möglicher Abweichungen wurden anschließend Spearman-Rangkorrelationen zwischen den korrespondierenden Dimensionskoordinaten berechnet. Unterschiede zwischen den dimensionsspezifischen Rangkorrelationen wurden mittels zweiseitiger Permutationstests ($n = 10.000$) geprüft, bei denen die Zuordnung der jeweils verglichenen

Dimensionen innerhalb der korrespondierenden Varianten permutiert wurde. Schließlich wurden die normierten INDSCAL-Gewichte deskriptiv gegenübergestellt, um mögliche Unterschiede in der relativen Bedeutung der einzelnen Dimensionen zwischen den Subgruppen abzubilden (Borg & Groenen, 2005; Carroll & Chang, 1970).

Zur Beantwortung von FF2 wurde untersucht, ob sich die Varianten auf Grundlage ihrer Positionen im dreidimensionalen Wahrnehmungsraum zu Gruppen zusammenfassen ließen und inwiefern die resultierenden Gruppierungen mit den musiktheoretisch definierten Kategorien übereinstimmten. Hierzu wurden mit dem R-Paket *cluster* (Version 2.1.8.1; Maechler et al., 2025) hierarchische Clusteranalysen auf Basis der euklidischen Distanzen zwischen den MDS-Koordinaten durchgeführt. Die Clusterbildung erfolgte mittels Average Linkage, bei dem der Abstand zwischen zwei Clustern als Mittelwert der paarweisen Distanzen ihrer Elemente bestimmt wurde. Zur Beurteilung der resultierenden Clusterbildung wurden mehrere Gütemaße herangezogen. Der cophenetische Korrelationskoeffizient (CCC) diente zur Beurteilung der strukturellen Passung der Clusterlösung, wobei Werte $> ,80$ als Hinweis auf eine gute Repräsentation der Distanzstruktur interpretiert werden (Everitt et al., 2011; Romesburg, 2004). Die Trennschärfe unterschiedlicher Clusterlösungen wurde anhand des mittleren Silhouettenkoeffizienten beurteilt, wobei mittlere Silhouettenkoeffizienten von $\bar{s} \geq ,51$ auf eine angemessene Clusterstruktur hinweisen (Kaufman & Rousseeuw, 1990). Ergänzend wurden mit dem R-Paket *mclust* (Version 6.1.2; Scrucca et al., 2023) Gaussian Mixture Models als modellbasierter Ansatz zur Identifikation möglicher Gruppierungen berechnet. Die Modellauswahl erfolgte anhand des Bayesian Information Criterion (BIC; Schwarz, 1978). Zur Quantifizierung der Übereinstimmung der Clusterzuordnungen zwischen Gesamtstichprobe und Subgruppen sowie zwischen den Subgruppen wurde der Adjusted Rand Index (ARI; Hubert & Arabie, 1985) herangezogen. Ein Wert von 1 entspricht einer vollständigen Übereinstimmung der Partitionen. Die resultierenden Gruppierungen wurden anschließend mit der musiktheoretischen Klassifikation der Varianten verglichen.

Ergebnisse

Prüfung der Vergleichbarkeit über die verschiedenen Tonfolgen

Die für jede Tonfolge auf Basis der vollständigen Stichprobe separat berechneten MDS-Konfigurationen wiesen durchweg sehr niedrige Stresswerte auf (3D-Stress: $M = ,02$; $SD = ,01$) und repräsentierten die Disparitäten exzellent (ausführliche Ergebnisse sind im Supplementmaterial auf OSF unter <https://osf.io/cfa83> verfügbar). Auch die Korrelationen der Disparitätsmatrizen zeigten mit durchschnittlichen Spearman- ρ -Werten im hohen Bereich ($\rho_M = ,86$; $SD = ,06$) eine starke Übereinstimmung der Distanzrangordnungen zwischen den Tonfolgen. Die Prokrustes-Korrelationen fielen im Mittel hoch aus ($r_M = ,93$; $SD = ,05$). Dies entsprach einer mittleren aufgeklärten Varianz von etwa

86% und wies auf eine ausgeprägte globale Übereinstimmung der Konfigurationen hin. Die permutationsbasierten PROTEST-Analysen ergaben für alle Vergleiche signifikante Übereinstimmungen (alle $p \leq ,003$). Die RMSE-Werte lagen im Mittel bei 0,31 ($SD = 0,05$). Dies entsprach etwa 15% der gesamten Skalenbreite und wies damit trotz hoher globaler Übereinstimmung auf moderate lokale Abweichungen zwischen den Konfigurationen hin. Insgesamt wiesen die Ergebnisse somit auf eine hohe Übereinstimmung der relativen Anordnung der Varianten über die Tonfolgen hinweg hin, bei gleichzeitig vorhandenen tonfolgenspezifischen metrischen Unterschieden.

Auch die konditionalen logistischen Regressionsmodelle deuteten insgesamt auf eine weitgehende Vergleichbarkeit der Ähnlichkeitsentscheidungen hin (Supplement D; <https://osf.io/cfa83>). Unterschiede zwischen den Variantenpaaren trugen wesentlich zur Modellgüte bei (McFaddens $\rho^2 = ,16$). Die zusätzliche Berücksichtigung tonfolgenspezifischer Unterschiede erhöhte die Modellgüte auf $\rho^2 = ,18$ ($\Delta\rho^2 = ,02$). Die Interaktion zwischen Variantenpaar und Tonfolge war im globalen clusterrobusten Wald-Test signifikant ($p < ,001$). Ein vergleichbares Muster zeigte sich in allen Subgruppen. Damit zeigte sich trotz der weitgehenden geometrischen Übereinstimmung eine geringe, jedoch statistisch signifikante Tonfolgenspezifität der Ähnlichkeitsentscheidungen. Die gemeinsame Modellierung erfolgte daher mittels INDSCAL, wodurch ein gemeinsamer Wahrnehmungsraum bei gleichzeitiger tonfolgenspezifischer Gewichtung der Dimensionen modelliert wurde.

Dimensionen der Ähnlichkeitswahrnehmung und Vergleich über Subgruppen (FF1)

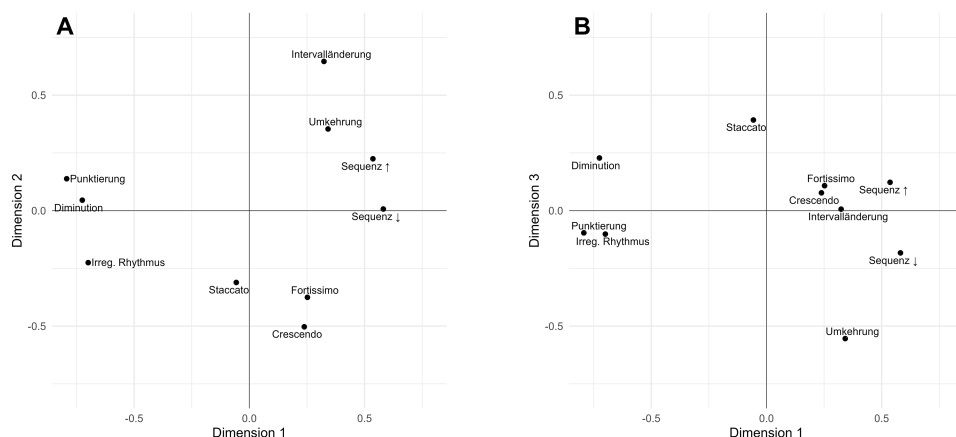
Dimensionalität und inhaltliche Charakterisierung

Im Hinblick auf FF1 zeigte der Vergleich von MDS-Lösungen mit zwei, drei und vier Dimensionen jeweils eine Verbesserung der Modellpassung mit zunehmender Dimensionenzahl, insbesondere von zwei auf drei Dimensionen (mittlere Stressreduktion = ,05), während der zusätzliche Zugewinn durch eine vierte Dimension vergleichsweise geringer ausfiel (mittlere Stressreduktion = ,02). Dies deutete darauf hin, dass die Ähnlichkeitsurteile durch einen dreidimensionalen Wahrnehmungsraum angemessen abgebildet wurden. Für die weiteren Analysen wurde daher eine dreidimensionale Lösung gewählt, die für die Gesamtstichprobe eine gute Modellpassung aufwies (Stress = ,07).

Die Anordnung der Varianten im dreidimensionalen Wahrnehmungsraum deutete auf eine inhaltlich interpretierbare Differenzierung entlang der ersten beiden Dimensionen hin ([Abbildung 3](#)).

Abbildung 3

Zweidimensionale Projektionen der dreidimensionalen MDS-Lösung des aggregierten Wahrnehmungsraums der gesamten Stichprobe



Anmerkung. Dargestellt sind die Positionen der zehn Varianten. Panel A: Dimension 1 × Dimension 2. Panel B: Dimension 1 × Dimension 3.

Die erste Dimension bildete eine deutliche Gegenüberstellung zwischen zeitlich-rhythmischer und tonhöhenbasierter Organisation ab, während die zweite Dimension zwischen performativ-klanglichen und relational-strukturellen Transformationen differenzierte. Die dritte Dimension ließ sich demgegenüber inhaltlich weniger eindeutig charakterisieren.

Die stärkste Differenzierung der Varianten zeigte sich entlang der ersten Dimension ($SD = 0,54$), gefolgt von der zweiten ($SD = 0,36$) und dritten Dimension ($SD = 0,26$). Die Subgruppen wichen in diesen Kennwerten nicht substantiell von der Gesamtstichprobe ab.

Die INDSCAL-Gewichte zeigten zugleich deskriptive Unterschiede in der tonfolgen-spezifischen Bedeutung der Dimensionen. Während Dimension 1 über die Tonfolgen hinweg vergleichsweise stabil gewichtet wurde ($Mdn = ,36$; $IQR = ,06$), zeigten die Gewichte der Dimensionen 2 ($Mdn = ,22$; $IQR = ,15$) und 3 ($Mdn = ,34$; $IQR = ,13$) deutlich stärkere tonfolgen-spezifische Unterschiede. Dies deutete darauf hin, dass die relative Bedeutung von Dimension 1 über die Tonfolgen hinweg vergleichsweise stabil war, während die Bedeutung der Dimensionen 2 und 3 stärker tonfolgen-spezifisch variierte.

Vergleichbarkeit der Subgruppen

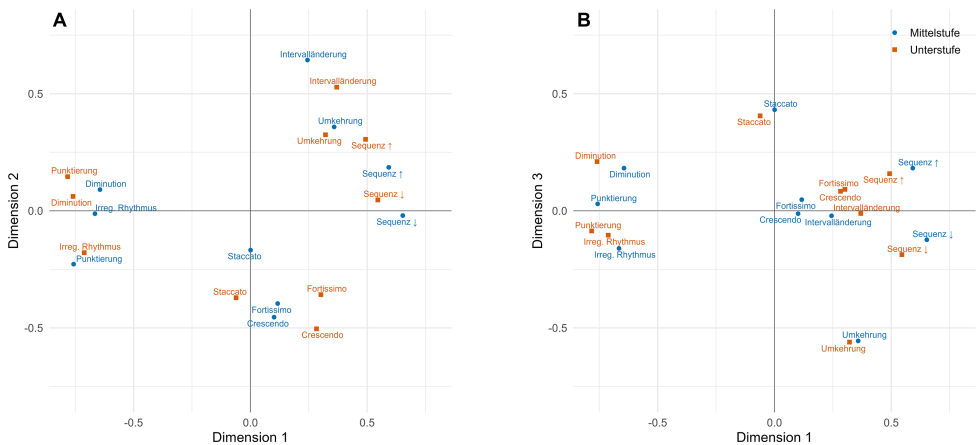
Die dreidimensionalen MDS-Lösungen wiesen auch in den Subgruppen gute bis akzeptable Modellpassungen auf (Stress = ,08 – ,10). Im Vergleich zur Gesamtlösung zeigten

sich hohe Prokrustes-Korrelationen ($r = ,92 - ,99$; alle PROTEST- $p = ,001$; Supplement E; <https://osf.io/cfa83>) bei RMSE-Werten zwischen 0,07 und 0,26. Insgesamt zeigte sich damit eine hohe globale Übereinstimmung der MDS-Konfigurationen über die Subgruppen hinweg.

Auch im direkten Vergleich der Subgruppen (siehe [Abbildungen 4 und 5](#)) zeigte sich eine hohe Übereinstimmung (Supplement E; <https://osf.io/cfa83>). Diese fiel zwischen Unter- und Mittelstufe höher aus (Prokrustes- $r = ,95$; RMSE = 0,19) als zwischen Instrumentalist:innen und Nichtinstrumentalist:innen (Prokrustes- $r = ,88$; RMSE = 0,29). Die dimensionsspezifischen Rangkorrelationen wurden ergänzend herangezogen, um die im globalen Vergleich beobachteten Abweichungen genauer zu lokalisieren (Supplement F; <https://osf.io/cfa83>). Dimensionsspezifisch zeigte sich für Unter- und Mittelstufe eine hohe Übereinstimmung. Zwischen Instrumentalist:innen und Nichtinstrumentalist:innen fiel die Übereinstimmung für die ersten beiden Dimensionen ebenfalls hoch aus (D1: Spearman- $\rho = ,96$; D2: $\rho = ,88$), während sie für die dritte Dimension deutlich geringer war (D3: $\rho = ,26$). Der Unterschied in der Übereinstimmung zwischen der dritten und ersten Dimension war statistisch signifikant ($\Delta\rho = -0,70$; $p = ,02$), nicht jedoch zwischen der dritten und zweiten Dimension ($\Delta\rho = -0,62$; $p = ,18$). Diese geringere Übereinstimmung der dritten Dimension ging bei den Instrumentalist:innen mit einer abweichenden Anordnung einzelner Varianten einher, die eine zusätzliche Differenzierung zwischen eher regelmäßig und eher irregulär wirkenden Varianten erkennen ließ.

Abbildung 4

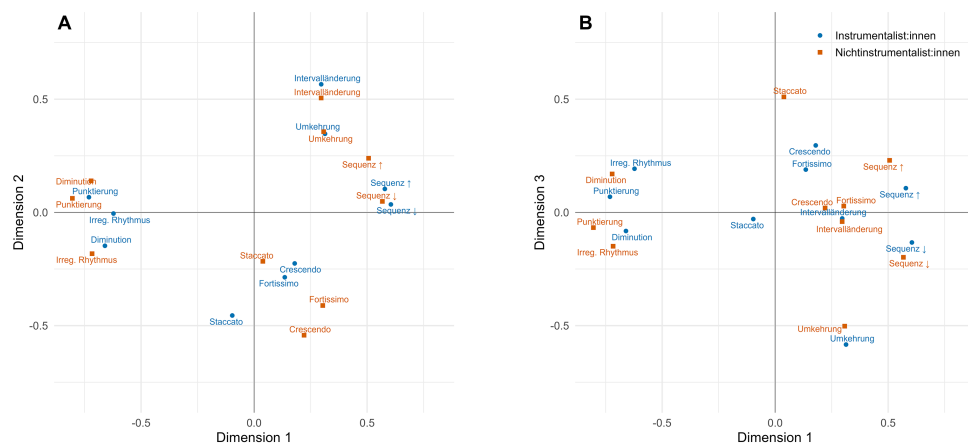
Zweidimensionale Projektionen der dreidimensionalen MDS-Lösung des aggregierten Wahrnehmungsraums im Vergleich Unter- zu Mittelstufe



Anmerkung. Dargestellt sind die Positionen der zehn Varianten. Panel A: Dimension 1 $\times$ Dimension 2. Panel B: Dimension 1 $\times$ Dimension 3.

Abbildung 5

Zweidimensionale Projektionen der dreidimensionalen MDS-Lösung des aggregierten Wahrnehmungsraums im Vergleich Instrumentalist:innen zu Nichtinstrumentalist:innen



Anmerkung. Dargestellt sind die Positionen der zehn Varianten. Panel A: Dimension 1 $\times$ Dimension 2. Panel B: Dimension 1 $\times$ Dimension 3.

Die INDSCAL-Gewichte zeigten deskriptiv über die Subgruppen hinweg eine weitgehend vergleichbare Gewichtung der Dimensionen (Supplement G; <https://osf.io/cfa83>). Dimension 2 wies in allen Subgruppen die geringsten medianen Gewichte auf, während die Dimensionen 1 und 3 tendenziell stärker gewichtet wurden. Insgesamt sprachen die Ergebnisse für einen dreidimensionalen Wahrnehmungsraum, dessen globale Anordnung über Schulstufen und Instrumentalstatus hinweg weitgehend vergleichbar war. Unterschiede zeigten sich zwischen Instrumentalist:innen und Nichtinstrumentalist:innen in der dritten, inhaltlich weniger eindeutig bestimmbar Dimension.

Gruppierung der Varianten (FF2)

Zur Identifikation möglicher Gruppierungen der Varianten ergaben hierarchische Clusteranalysen (Supplement H; <https://osf.io/cfa83>) über die Gesamtstichprobe und die Subgruppen hinweg hohe cophenetische Korrelationskoeffizienten (Gesamtstichprobe: $r = ,88$; Subgruppen: $r_M = ,89$; $SD = ,04$). Die hohen Koeffizienten zeigten, dass die Dendrogramme die zugrunde liegenden Distanzen des dreidimensionalen Wahrnehmungsraums gut bewahrten. Die höchsten Silhouettenkoeffizienten ergaben sowohl für die Gesamtstichprobe als auch für alle Subgruppen eine konsistente Dreiclusterlösung ($k = 3$). Die Silhouettenwerte lagen zwischen ,47 und ,54 ($\bar{s}_M = ,51$; $SD = ,02$). Während in allen Gruppen mit Ausnahme der Mittelstufe Werte bei mindestens $\bar{s} = ,51$ auf eine angemessene Clusterstruktur hinwiesen, deutete der Wert für die Mittelstufe mit $\bar{s} = ,47$

auf eine etwas schwächere, aber weiterhin vorhandene Dreiclusterstruktur hin. Die Gaussian Mixture Models stützten diese Struktur ebenfalls, da in sämtlichen Subgruppen die beste Modellanpassung für $k = 3$ erzielt wurde. Für die paarweisen Vergleiche der Clusterzuordnungen zwischen den Stichproben betrug der Adjusted Rand Index (ARI) bei $k = 3$ durchgängig 1,00. Damit stimmten die aus den hierarchischen Clusteranalysen hervorgegangenen Partitionen sowohl zwischen den Teilstichproben als auch im Vergleich zur Gesamtstichprobe vollständig überein, wobei die Varianten jeweils dieselbe Dreiclusterpartition bildeten. Die resultierende Dreiclusterlösung und die Zuordnung der Varianten sind in [Tabelle 3](#) dargestellt. Die empirisch ermittelten Gruppierungen wurden anschließend mit der musiktheoretischen Klassifikation der Varianten verglichen. Dabei zeigte sich eine inhaltliche Entsprechung: Cluster 1 umfasste die oberflächenbezogenen Varianten, Cluster 2 die tonhöhenbasierten Varianten und Cluster 3 die zeitlich-rhythmischen Varianten. Die empirische Gruppierung spiegelte damit die musiktheoretische Klassifikation der Varianten wider.

Tabelle 3
Zuordnung der Varianten zu den identifizierten Clustern

Cluster 1	Cluster 2	Cluster 3
<ul style="list-style-type: none">• Fortissimo• Crescendo• Staccato	<ul style="list-style-type: none">• Sequenz ↑• Sequenz ↓• Umkehrung• Intervalländerung	<ul style="list-style-type: none">• Diminution• Irregulärer Rhythmus• Punktierung

Anmerkung. Die Tabelle zeigt die inhaltliche Zuordnung der musikalischen Varianten zu den drei identifizierten Clustern auf Grundlage der hierarchischen Clusteranalyse und der Gaussian Mixture Models.

Diskussion

Die vorliegende Studie zielte darauf ab, einen konsistenten Wahrnehmungsraum melodischer Variantenarten bei Schüler:innen zu identifizieren, dessen Vergleichbarkeit zwischen Gruppen unterschiedlicher musikalischer Expertise zu prüfen und die resultierenden Gruppierungen hinsichtlich ihrer Übereinstimmung mit musiktheoretischen Klassifikationen einzuordnen. Im Hinblick auf die Frage, ob sich auch unter Berücksichtigung von musikalischen Expertiseunterschieden ein konsistenter Wahrnehmungsraum finden lässt, sprechen die Befunde zunächst für einen stabilen dreidimensionalen Wahrnehmungsraum melodischer Variantenarten. Die erste und zugleich stabilste Dimensionsachse verweist dabei auf eine grundlegende Unterscheidung zwischen zeitlich-rhythmischer und tonhöhenbasierter Organisation melodischer Variantenarten. Damit scheint für die Wahrnehmung melodischer Ähnlichkeit insbesondere die Unterscheidung zwischen zeitlich-rhythmischen und tonhöhenbasierten Strukturmerkmalen zentral zu

sein. Dieses Ergebnis steht im Einklang mit empirischen Befunden, nach denen zeitliche und tonhöhenbasierte Parameter zentrale und teilweise differenziert verarbeitete Organisationsdimensionen musikalischer Wahrnehmung darstellen (Lappe et al., 2013, 2016; Prince, 2014).

Entlang der zweiten Dimension zeigt sich insbesondere eine Gegenüberstellung dynamischer und artikulatorischer Veränderungen einerseits und stärker relational-struktureller Transformationen, etwa intervall- und sequenzbezogener Variantenarten, andererseits. Diese Anordnung könnte auf Unterschiede zwischen Veränderungen der klanglichen Realisierung und Eingriffen in relationale Strukturmerkmale hinweisen. Zudem sind diese Unterschiede anschlussfähig an Ansätze hierarchischer Musikverarbeitung, nach denen elementare akustische Eigenschaften früh verarbeitet, während komplexere strukturelle Relationen erst in späteren, aufmerksamkeitsabhängigen Stufen integriert werden (Brattico et al., 2018; Jäncke, 2024; Koelsch, 2011; Koelsch & Schröger, 2018).

Die dritte Dimension verweist demgegenüber weniger auf eine klar umrissene eigenständige Organisationslogik als vielmehr auf residuale und stärker tonfolgenabhängige perzeptive Einflüsse. Die vergleichsweise geringe Streuung der Varianten entlang dieser Achse könnte darauf hindeuten, dass sich die Ähnlichkeitsurteile hier weniger deutlich voneinander unterscheiden. Die zusätzliche Reduktion des Stresswertes beim Übergang zur vierdimensionalen Lösung könnte auch darauf hindeuten, dass über die ersten beiden dominanten Dimensionsachsen hinaus weitere, inhaltlich weniger eindeutig bündelbare Einflussfaktoren wirksam sind, die sich in der dritten Dimension überlagern. Zugleich weist die hohe Streuung der INDSCAL-Gewichte darauf hin, dass die relative Bedeutung dieser Dimension zwischen den Tonfolgen stärker variiert. Dies könnte auf tonfolgen-spezifische Kontexteinflüsse zurückzuführen sein und steht im Einklang mit bisherigen Studien, die die Bedeutung des Kontexts für Ähnlichkeitsurteile hervorheben (Ahlbäck, 2007; Cambouropoulos, 2009; Taher et al., 2018). Hinsichtlich expertisearchhängiger Unterschiede zeigt sich die identifizierte Wahrnehmungsstruktur als weitgehend stabil. Sowohl zwischen Unter- und Mittelstufe als auch zwischen Instrumentalist:innen und Nichtinstrumentalist:innen blieben die ersten beiden Dimensionsachsen in ihrer grundlegenden Ausrichtung weitgehend erhalten.

Unterschiede in Abhängigkeit vom Expertisierungsgrad deuteten sich vor allem entlang der dritten Dimension an. Die ausbleibenden Unterschiede auf der zweiten Dimensionsachse sind dabei vor dem Hintergrund früherer Studien zu betrachten, die bei musikalisch expertisierteren Personen eine stärkere Sensitivität gegenüber komplexeren strukturellen Relationen nahelegen (Bonetti et al., 2022; Eitan & Granot, 2009). Eine mögliche Erklärung könnte darin liegen, dass auch Nichtinstrumentalist:innen bereits über grundlegende, durch implizites musikalisches Lernen erworbene Strukturierungsstrategien verfügen, die für die Verarbeitung vergleichsweise kurzer Tonfolgen weitgehend ausreichen. Zusätzliche, durch Instrumentalunterricht erworbene Differenzierungsvorteile

würden unter diesen relativ einfachen Wahrnehmungsbedingungen daher nur begrenzt sichtbar.

Während sich bei den Nichtinstrumentalist:innen in dieser Dimension keine klar abgrenzbare inhaltliche Ordnung erkennen lässt, zeigen sich bei den Instrumentalist:innen Hinweise auf eine Differenzierung zwischen eher in sich regelhaft aufgebauten und eher unregelmäßigen Variantenarten. Diese Interpretation ist anschlussfähig an Befunde, wonach musikalische Expertise mit differenzierteren auditiven Erwartungsmodellen sowie einer erhöhten Sensitivität gegenüber Regelverletzungen einhergeht (Bonetti et al., 2022, 2024; van Zuijen et al., 2004, 2005; Zhao et al., 2017). In Bezug auf den Expertisierungsgrad deuten diese Befunde insgesamt darauf hin, dass unter den hier untersuchten Stimulusbedingungen außerunterrichtlicher Instrumental- oder Vokalunterricht weniger mit Unterschieden in den grundlegenden Dimensionsachsen melodischer Ähnlichkeitswahrnehmung als mit subtileren Unterschieden in der Gewichtung einzelner Varianten verbunden ist.

Zwischen Unter- und Mittelstufe zeigten sich auf allen drei Dimensionsachsen keine strukturellen Unterschiede. Besonders hervorzuheben ist dieser Befund vor dem Hintergrund entwicklungspsychologischer Ansätze, die eine altersbedingte Zunahme der Arbeitsgedächtniskapazität und des Abstraktionsvermögens nahelegen (Cowan, 2016; Jentschke & Koelsch, 2006), was die Integration größerer Informationsmengen sowie differenziertere strukturelle Analysen begünstigen kann. Auch der Musikunterricht im baden-württembergischen Bildungsplan ist spiralcurricular angelegt und sieht eine sukzessive Ausdifferenzierung musiktheoretischer Inhalte vor (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016). Das Ausbleiben entsprechender Unterschiede könnte darauf hindeuten, dass die für die vorliegenden Stimuli und die Aufgabenkomplexität relevanten Wahrnehmungs- und Organisationsprozesse bereits in der Unterstufe weitgehend ausgebildet sind (Degé & Roden, 2018; Thompson, 2009). Zugleich könnte dies darauf hindeuten, dass die im schulischen Verlauf zunehmende Auseinandersetzung mit musiktheoretischen Inhalten nicht unmittelbar mit Veränderungen in der Organisation von Ähnlichkeitsurteilen einhergeht. Eine mögliche Erklärung könnte darin liegen, dass explizites musiktheoretisches Wissen im schulischen Musikunterricht nur begrenzt mit mental-auditiven Repräsentationen verknüpft wird (Ruf, 2024) und sich eine zunehmende musiktheoretische Differenzierung daher nicht unmittelbar in der Organisation wahrgenommener Ähnlichkeit widerspiegelt.

Im Hinblick auf die Frage, ob sich aus dem gefundenen Wahrnehmungsraum musiktheoretisch begründete Cluster der Varianten bilden lassen, legt die über verschiedene Verfahren hinweg gestützte Dreiclusterlösung nahe, dass sich musiktheoretische Differenzierungen melodischer Variantenarten auch in der Wahrnehmung der Schüler:innen abbilden. Die identifizierten Cluster lassen sich oberflächenbezogenen, tonhöhenbasierten sowie rhythmisch-zeitlichen Variantenarten zuordnen. Die Ergebnisse weisen damit auf eine Übereinstimmung zwischen den empirisch ermittelten Gruppierungen und den

musiktheoretischen Klassifikationen hin, wobei die Variantenarten zugleich innerhalb eines mehrdimensionalen Wahrnehmungsraums angeordnet sind.

Limitationen

Bei der Interpretation der Ergebnisse sind einige Limitationen zu berücksichtigen, die insbesondere die Gestaltung der Stimuli, die Auswertung der Ähnlichkeitsurteile sowie die Interpretation des resultierenden Wahrnehmungsraums betreffen.

Eine Einschränkung betrifft die Auswahl und Gestaltung der Stimuli. Die verwendeten Tonfolgen waren bewusst kurz und vergleichsweise einfach gehalten und deckten hinsichtlich des Tonmaterials sowie der Ausprägung der Varianten nur einen begrenzten Ausschnitt möglicher musikalischer Gestaltungen ab. Die Varianten wurden zwar systematisch für alle Tonfolgen erstellt, repräsentierten jedoch jeweils einzelne musikalische Merkmale. Kombinationen mehrerer Variantenarten wurden nicht untersucht. Der resultierende Wahrnehmungsraum ist daher zunächst auf diesen Stimulusbereich bezogen und bildet möglicherweise nicht die gesamte Differenziertheit ab, die sich bei einem breiteren oder komplexeren Stimulusmaterial ergeben könnte. Dies ist vermutlich insbesondere für die zweite und dritte Dimension relevant, entlang der die Varianten eine geringere Streuung aufweisen. Auch die ausbleibenden Unterschiede zwischen Unter- und Mittelstufe sind vor diesem Hintergrund vorsichtig zu interpretieren. Es ist nicht auszuschließen, dass die begrenzte Komplexität der Tonfolgen sowie die moderate Ausprägung und Spannbreite der Varianten feinere Differenzierungen des Wahrnehmungsraums und mögliche schulstufenbezogene Unterschiede weniger deutlich hervortreten ließen. Die Begrenzung der Stimuluskomplexität erfolgte bewusst mit Blick auf die Zielgruppe und die Durchführbarkeit des Tripleparadigmas. Die vorliegenden Ergebnisse sind daher als Ausgangspunkt für eine weiterführende Prüfung des Wahrnehmungsraums unter erweiterten Stimulusbedingungen zu verstehen. Zukünftige Studien sollten gezielt prüfen, ob sich die identifizierten Dimensionen bei längeren und komplexeren Tonfolgen sowie einer größeren Spannbreite und stärkeren Ausprägung der Varianten deutlicher differenzieren und inwiefern sich die Befunde unter diesen Bedingungen bestätigen.

Eine weitere Einschränkung betrifft die Aggregation der im Tripleparadigma erhobenen Ähnlichkeitsentscheidungen zu Proximitäten. Zwar wurde durch die ausgewogene Präsentation der Triple-Konstellationen sichergestellt, dass unterschiedliche Vergleichskontexte gleichgewichtet in die resultierenden Proximitäten eingingen. Durch die Aggregation bildet die MDS jedoch ausschließlich die über diese Kontexte hinweg resultierende globale Ähnlichkeitsstruktur ab. Sie erlaubt daher keine Aussage darüber, ob und inwiefern einzelne Paarentscheidungen von der jeweils dritten präsentierten Variante abhingen. Die vorliegenden Ergebnisse beziehen sich somit auf die über die Vergleichskontexte aggregierte Wahrnehmungsstruktur der Variantenarten. Zukünftige Untersuchungen sollten deshalb die Modellierung des Wahrnehmungsraums erweitern, indem die jeweils dritte Variante explizit berücksichtigt wird, um neben der globalen

Ähnlichkeitsstruktur auch mögliche kontextabhängige Unterschiede in den Ähnlichkeitsurteilen und der relativen Salienz der Varianten abzubilden.

Eine weitere Einschränkung betrifft die inhaltliche Interpretation der Dimensionen des Wahrnehmungsraums. Die mittels MDS identifizierten Achsen sind nicht als musikalische Merkmalsdimensionen im musiktheoretischen Sinne zu verstehen, sondern ergeben sich aus der räumlichen Repräsentation der Ähnlichkeitsurteile. Entsprechend können sich entlang einer Achse Variantenarten gegenüberstehen, die musiktheoretisch unabhängige bzw. orthogonale Merkmale betreffen, wie etwa Rhythmus und Tonhöhe. Eine solche Gegenüberstellung ist daher nicht als musikalisches Kontinuum zwischen Rhythmus und Tonhöhe zu interpretieren, sondern als empirisch resultierende Trennung innerhalb der Ähnlichkeitsurteile der Schüler:innen. Der Erkenntniswert des dreidimensionalen Wahrnehmungsraums liegt dabei über der Differenzierung einzelner Variantenarten hinaus insbesondere in deren relationaler Einordnung zueinander. Die Untersuchung zielte entsprechend auf die Beschreibung der wahrnehmungsbezogenen Organisation unterschiedlicher Variantenarten und nicht auf die Rekonstruktion eines musiktheoretischen Merkmalsraums. Inwiefern sich innerhalb dieser empirischen Struktur musiktheoretisch interpretierbare Gruppierungen zeigten, wurde ergänzend anhand der Clusterstruktur untersucht. Gerade diese wahrnehmungsbezogene Perspektive erscheint für den schulischen Kontext relevant, da sie Hinweise darauf geben kann, wie Schüler:innen unterschiedliche Varianten wahrnehmungsseitig zueinander einordnen und an welchen Stellen weiterführende diagnostische oder förderbezogene Untersuchungen ansetzen könnten.

Schlussfolgerung

Die Befunde legen nahe, dass die Wahrnehmung melodischer Ähnlichkeit auf mehreren Dimensionen beruht, die sowohl stabile als auch tonfolgen- und expertiseabhängige Unterschiede aufweisen. Vor dem Hintergrund des hier untersuchten Alters, in dem insbesondere die Schule einen zentralen Rahmen für formales musikalisches Lernen darstellt, kommt diesem Befund besondere Bedeutung zu. Die Ergebnisse deuten darauf hin, dass musikalische Parameter, wie sie im schulischen Musikunterricht thematisiert werden, nicht losgelöst von ihrer wahrnehmungsbezogenen Organisation betrachtet werden sollten. Zugleich sprechen die Befunde dafür, dass sich zentrale musiktheoretische Differenzierungen in der wahrnehmungsbezogenen Organisation der untersuchten Schüler:innen widerspiegeln und innerhalb eines mehrdimensionalen Wahrnehmungsraums relational eingebettet sind. Eine stärkere Berücksichtigung impliziter Strukturwahrnehmung könnte dazu beitragen, theoretische Konzepte systematisch an bestehende perzeptive Organisationsformen im Kindes- und Jugendalter anzubinden. Für das Verständnis musikalischen Lernens in diesem Altersbereich ergibt sich daraus die Perspektive, gezielter zu erfassen, entlang welcher Organisationsdimensionen Kinder und Jugendliche melodische Varianten differenzieren und in welchen Bereichen sich individuelle Unterschiede zeigen.

Finanzierung: Die Autor*innen haben keine Finanzierung für das Forschungsprojekt erhalten.

Danksagung: Wir bedanken uns herzlich bei Jonas Zarzalis für die Unterstützung der Implementierung des Fragebogens auf SoSci Survey sowie für die Hilfe bei der Datenaufbereitung.

Interessenkonflikte: Die Autor*innen haben erklärt, dass keinerlei konkurrierende Interessen bestehen.

Beiträge der Autor*innen: Myrto Zarzalis – Konzeption | Datenkuratierung | Formale Analyse | Untersuchung | Visualisierung | Methodik | Projektadministration | Ressourcen | Software | Verfassen – Originalentwurf | Schreiben – Überarbeitung & Redaktion.

Johannes Hasselhorn – Methodik | Supervision | Schreiben – Überarbeitung & Redaktion.

Daniel Fiedler – Supervision | Schreiben – Überarbeitung & Redaktion.

Ethikerklärung: Die Studie wurde in Übereinstimmung mit den einschlägigen institutionellen und nationalen Richtlinien und Vorschriften (Föderation Deutscher Psychologinnenvereinigungen, 2022) sowie im Einklang mit den Prinzipien der Deklaration von Helsinki durchgeführt. Eine formale Genehmigung der Studie durch eine Ethikkommission war nicht erforderlich. Alle Schüler:innen sowie ihre Sorgeberechtigten gaben ihre schriftliche Einwilligung und wurden darüber informiert, dass sie das Recht haben, ihre Teilnahme an der Studie jederzeit ohne negative Konsequenzen abzubrechen. Die Studie wurde durch das Ministerium für Kultus, Jugend und Sport Baden-Württemberg genehmigt.

KI Nutzung: Für die sprachliche Überarbeitung und die Korrektur von Grammatik und Rechtschreibung dieses Manuskripts wurde GPT-OSS 120B (Stand: April/2026) verwendet. Für die Unterstützung bei der Literaturrecherche wurden Elicit (Web-App, Stand: April/2026) und ResearchRabbit (Web-App, Stand: April/2026) verwendet. Bei der Erstellung und Optimierung der Auswertungsskripte in R wurden unterstützend GPT-OSS 120B (Stand: April/2026) und Qwen (Version Qwen3.5 397B A17B) eingesetzt. Die endgültige inhaltliche Kontrolle sowie die Überprüfung der wissenschaftlichen Korrektheit erfolgten durch die Autor:innen.

Datenverfügbarkeit: Aus datenschutzrechtlichen Gründen ist die Nutzung der Rohdaten den Autor:innen vorbehalten. Die aggregierten Daten, vollständige Ergebnisse der einzelnen Analysen sowie der verwendete Fragebogen sind über das Open Science Framework (OSF) unter <https://osf.io/cfa83> einsehbar.

Ergänzende Materialien

Die aggregierten Daten, vollständige Ergebnisse der einzelnen Analysen sowie der verwendete Fragebogen sind über das Open Science Framework (OSF) einsehbar (siehe Zarzalis, 2026).

Quellenverzeichnis der ergänzenden Materialien

Zarzalis, M. (2026). *Wahrnehmungsräume melodischer Varianten* [Daten, Materialien]. OSF.
<https://osf.io/cfa83>

Literatur

- Ahlbäck, S. (2007). Melodic similarity as a determinant of melody structure. *Musicae Scientiae*, 11(1, Suppl.), 235–280. <https://doi.org/10.1177/102986490701100110>
- Altenmüller, E. (2018). Hirnphysiologie musikalischen Lernens. In W. Gruhn & P. Rübke (Hrsg.), *Musiklernen: Bedingungen – Handlungsfelder – Positionen* (S. 43–69). Helbling.
- Bonetti, L., Brattico, E., Carlomagno, F., Cabral, J., Stevner, A., Deco, G., Whybrow, P. C., Pearce, M., Pantazis, D., Vuust, P., & Kringelbach, M. L. (2024). Spatiotemporal whole-brain activity and functional connectivity of melodies recognition. *Cerebral Cortex*, 34(8), Article bhae320. <https://doi.org/10.1093/cercor/bhae320>
- Bonetti, L., Carlomagno, F., Kluchko, M., Gold, B. P., Palva, S., Haumann, N. T., Tervaniemi, M., Huotilainen, M., Vuust, P., & Brattico, E. (2022). Whole-brain computation of cognitive versus acoustic errors in music: A mismatch negativity study. *Neuroimage. Reports*, 2(4), Article 100145. <https://doi.org/10.1016/j.ynirp.2022.100145>
- Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications* (2. Aufl.). Springer.
- Borg, I., Groenen, P. J. F., & Mair, P. (2018). *Applied multidimensional scaling and unfolding*. Springer. <https://doi.org/10.1007/978-3-319-73471-2>
- Brattico, E., Olcese, C., & Tervaniemi, M. (2018). Automatic processing of musical sounds in the human brain. In R. Bader (Hrsg.), *Springer handbook of systematic musicology* (S. 441–452). Springer. https://doi.org/10.1007/978-3-662-55004-5_22
- Cambouropoulos, E. (2009). How similar is similar? *Musicae Scientiae*, 13(1, suppl.), 7–24. <https://doi.org/10.1177/102986490901300102>
- Cao, E., Lotstein, M., & Johnson-Laird, P. N. (2014). Similarity and families of musical rhythms. *Music Perception*, 31(5), 444–469. <https://doi.org/10.1525/mp.2014.31.5.444>
- Carroll, J. D., & Chang, J.-J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of “Eckart-Young” decomposition. *Psychometrika*, 35(3), 283–319. <https://doi.org/10.1007/BF02310791>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Routledge. <https://doi.org/10.4324/9780203771587>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, 24(1), 87–114. <https://doi.org/10.1017/S0140525X01003922>
- Cowan, N. (2016). *Working memory capacity*. Routledge.
- Degé, F., & Roden, I. (2018). Entwicklung musikalischer Fähigkeiten: Kindergarten und Grundschule. In A. C. Lehmann & R. Kopiez (Hrsg.), *Handbuch Musikpsychologie* (S. 151–180). Hogrefe.
- Deliège, I. (2001). Similarity perception ↔ categorization ↔ cue abstraction. *Music Perception*, 18(3), 233–243. <https://doi.org/10.1525/mp.2001.18.3.233>
- Dress, S., & von Fischer, K. (2016). Variation: Einführung, Systematisches. In L. Lütteken (Hrsg.), *MGG Online*. RILM; Bärenreiter; Metzler. <https://www.mgg-online.com/mgg/stable/596596>

- Eerola, T., & Bregman, M. (2007). Melodic and contextual similarity of folk song phrases. *Musicae Scientiae*, 11(1, Suppl.), 211–233. <https://doi.org/10.1177/102986490701100109>
- Eitan, Z., & Granot, R. Y. (2009). Primary versus secondary musical parameters and the classification of melodic motives. *Musicae Scientiae*, 13(1, Suppl.), 139–179. <https://doi.org/10.1177/102986490901300107>
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. Wiley. <https://doi.org/10.1002/9780470977811>
- Föderation Deutscher Psychologenvereinigungen. (2022). *Berufsethische Richtlinien*. https://www.dgps.de/fileadmin/user_upload/PDF/Berufsethische_Richtlinien/BER-Foederation-20230426-Web-1.pdf
- Frieler, K. (2018). Gruppierung, Ordnung und Ähnlichkeit in der Musik. In A. C. Lehmann & R. Kopiez (Hrsg.), *Handbuch Musikpsychologie* (S. 513–544). Hogrefe.
- Götte, U. (2020). *Wiederholung als zentrales universelles Gestaltungsmittel der Musik*. J.B. Metzler. <https://doi.org/10.1007/978-3-476-05635-1>
- Gruber, H., & Lehmann, A. C. (2014). Begabung, Talent und Expertise. In M. Stamm (Hrsg.), *Handbuch Talententwicklung: Theorien, Methoden und Praxis in Psychologie und Pädagogik* (1. Aufl., S. 349–364). Hans Huber.
- Hahn, U. (2014). Similarity. *Wiley Interdisciplinary Reviews. Cognitive Science*, 5(3), 271–280. <https://doi.org/10.1002/wcs.1282>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8. Aufl.). Cengage.
- Harrison, P. M., Musil, J. J., & Müllensiefen, D. (2016). Modelling melodic discrimination tests: Descriptive and explanatory approaches. *Journal of New Music Research*, 45(3), 265–280. <https://doi.org/10.1080/09298215.2016.1197953>
- Hasselhorn, J., & Knigge, J. (2026). Kompetenz und Expertise. In M. Dartsch, J. Knigge, A. Niessen, F. Platz, & C. Stöger (Hrsg.), *Handbuch Musikpädagogik: Grundlagen – Forschung – Diskurse* (2. Aufl., S. 385–398). Waxmann.
- Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1), 193–218. <https://doi.org/10.1007/BF01908075>
- Jäncke, L. (2024). *Lehrbuch kognitive Neurowissenschaften* (4. Aufl.). Hogrefe.
- Jentschke, S., & Koelsch, S. (2006). Gehirn, Musik, Plastizität und Entwicklung. In A. Scheunpflug & C. Wulf (Hrsg.), *Zeitschrift für Erziehungswissenschaft* (S. 51–70). https://doi.org/10.1007/978-3-531-90607-2_5
- Jungmann, B., Schauf, C., Lemmel, H., Herden, J., Hörnel, D., & Hübenthal, M. (2024). *Capella* (Version 10.0.6.0) [Computer software]. capella-software AG. <https://www.capella-software.com/>
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. Wiley.
- Koelsch, S. (2011). Toward a neural basis of music perception: A review and updated model. *Frontiers in Psychology*, 2, Article 110. <https://doi.org/10.3389/fpsyg.2011.00110>

- Koelsch, S., & Schröger, E. (2018). Neurowissenschaftliche Grundlagen der Musikverarbeitung. In A. C. Lehmann & R. Kopiez (Hrsg.), *Handbuch Musikpsychologie* (S. 461–481). Hogrefe.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11(3), 346–374. [https://doi.org/10.1016/0010-0285\(79\)90016-1](https://doi.org/10.1016/0010-0285(79)90016-1)
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195148367.001.0001>
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1), 1–27. <https://doi.org/10.1007/BF02289565>
- Kühn, C. (2021). Form. In L. Lütteken (Hrsg.), *MGG Online. RILM*; Bärenreiter; Metzler. <https://www.mgg-online.com/mgg/stable/534814>
- Lamont, A., & Dibben, N. (2001). Motivic structure and the perception of similarity. *Music Perception*, 18(3), 245–274. <https://doi.org/10.1525/mp.2001.18.3.245>
- Lappe, C., Lappe, M., & Pantev, C. (2016). Differential processing of melodic, rhythmic and simple tone deviations in musicians - an MEG study. *NeuroImage*, 124, 898–905. <https://doi.org/10.1016/j.neuroimage.2015.09.059>
- Lappe, C., Steinsträter, O., & Pantev, C. (2013). Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection. *Frontiers in Human Neuroscience*, 7, Article 260. <https://doi.org/10.3389/fnhum.2013.00260>
- Legendre, P. (2012). *Numerical ecology* (3. Aufl.). Elsevier.
- Leiner, D. J. (2024). *SoSci Survey* (Version 3.5.02) [Computer software]. <https://www.sosicisurvey.de>
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2025). *cluster: Cluster analysis basics and extensions* (Version 2.1.8.1) [R package]. <https://CRAN.R-project.org/package=cluster>
- Mair, P., Groenen, P. J. F., & de Leeuw, J. (2022). More on multidimensional scaling and unfolding in R: Smacof Version 2. *Journal of Statistical Software*, 102(10), 1–47. <https://doi.org/10.18637/jss.v102.i10>
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Hrsg.), *Economic theory and mathematical economics. Frontiers in econometrics* (S. 105–142). Academic Press.
- Meyer, T. (2024). Von Wellenformen, schleichenden Löwen und ausgerutschten Tönen. *Beiträge empirischer Musikpädagogik*, 15, 1–13. <https://doi.org/10.62563/bem.v2024236>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <https://doi.org/10.1037/h0043158>
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2016). *Bildungsplan des Gymnasiums: Musik*. https://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_MUS.pdf
- Müllensiefen, D., & Frieler, K. (2007). Modelling experts' notions of melodic similarity. *Musicae Scientiae*, 11(1, Suppl.), 183–210. <https://doi.org/10.1177/102986490701100108>

- Novello, A., McKinney, M. M., & Kohlrausch, A. (2011). Perceptual evaluation of inter-song similarity in western popular music. *Journal of New Music Research*, 40(1), 1–26.
<https://doi.org/10.1080/09298215.2010.523470>
- Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Borman, T., Carvalho, G., Chirico, M., de Caceres, M., . . . Weedon, J. (2025). *vegan: Community ecology package* (Version 2.7-2) [R package].
<https://CRAN.R-project.org/package=vegan>
- Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: A meta-analysis. *Frontiers in Psychology*, 5, Article 646.
<https://doi.org/10.3389/fpsyg.2014.00646>
- Prince, J. B. (2014). Contributions of pitch contour, tonality, rhythm, and meter to melodic similarity. *Journal of Experimental Psychology. Human Perception and Performance*, 40(6), 2319–2337. <https://doi.org/10.1037/a0038010>
- R Core Team. (2025). *R: A language and environment for statistical computing* (Version 4.5.2) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org>
- Rohrmeier, M., & Rebuschat, P. (2012). Implicit learning and acquisition of music. *Topics in Cognitive Science*, 4(4), 525–553. <https://doi.org/10.1111/j.1756-8765.2012.01223.x>
- Romesburg, H. C. (2004). *Cluster analysis for researchers*. Lulu.
- Ruf, B. (2024). *Musiktheorie im Musikunterricht: Eine qualitative Studie zum Deutungswissen von Lehrenden*. Allitera.
- Schönberg, A. (1979). *Die Grundlagen der musikalischen Komposition*. Universal-Edition.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461–464.
<https://doi.org/10.1214/aos/1176344136>
- Scrucca, L., Fraley, C., Murphy, T. B., & Raftery, A. E. (2023). *Model-based clustering, classification, and density estimation using mclust in R*. Chapman and Hall/CRC.
<https://doi.org/10.1201/9781003277965>
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27(2), 125–140. <https://doi.org/10.1007/BF02289630>
- Struktur (allgemein). (o. D.). In *Brockhaus Enzyklopädie Online*. Abgerufen am 16.04.2026 von <https://brockhaus.de/ecs/enzy/article/struktur-allgemein>
- Taher, C., Hasegawa, R., & McAdams, S. (2018). Effects of musical context on the recognition of musical motives during listening. *Music Perception*, 36(1), 77–97.
<https://doi.org/10.1525/mp.2018.36.1.77>
- Tenzer, M. (Hrsg.). (2006). *Analytical studies in world music*. Oxford University Press.
- Therneau, T. M. (2026). *A package for survival analysis in R* (Version 3.8-9) [R package].
<https://cran.r-project.org/web/packages/survival/index.html>
- Therneau, T. M., & Grambsch, P. M. (2000). *Modeling survival data: Extending the Cox model*. Springer.

- Thompson, W. F. (2009). *Music, thought, and feeling: Understanding the psychology of music*. Oxford University Press.
- Train, K. E. (2009). *Discrete choice methods with simulation* (2. Aufl.). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511805271>
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327–352.
<https://doi.org/10.1037/0033-295X.84.4.327>
- van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds: An event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience*, 16(2), 331–338. <https://doi.org/10.1162/089892904322984607>
- van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity: A mismatch negativity study comparing musicians and non-musicians. *Brain Research. Cognitive Brain Research*, 23(2–3), 270–276. <https://doi.org/10.1016/j.cogbrainres.2004.10.007>
- Zarzalís, M., Fiedler, D., Zarzalís, J., & Hasselhorn, J. (2026). Melodische Tonfolgen im schulischen Musikunterricht: Eine korpusbasierte Analyse von Musikstücken und -werken verschiedener Epochen, Genres und Gattungen. *Beiträge empirischer Musikpädagogik*, 18(2), 1–23.
<https://doi.org/10.62563/bem.v18i2.286>
- Zhao, T. C., Lam, H. T. G., Sohi, H., & Kuhl, P. K. (2017). Neural processing of musical meter in musicians and non-musicians. *Neuropsychologia*, 106, 289–297.
<https://doi.org/10.1016/j.neuropsychologia.2017.10.007>



Das *Jahrbuch Musikpsychologie* (JBDGM) ist die offizielle Zeitschrift der Deutschen Gesellschaft für Musikpsychologie (DGM).



PsychOpen GOLD ist ein Publikationsdienst des Leibniz-Instituts für Psychologie (ZPID), Deutschland.