

Freie Forschungsberichte

Urteilshomogenität und Klassengemeinschaft – Ein Beitrag zur Offenohrigkeitshypothese

Edda Leopold

Zusammenfassung

Die Offenohrigkeitshypothese nimmt an, dass jüngere Kinder offener gegenüber unkonventioneller Musik sind als ältere. Dieser Beitrag ist motiviert durch zwei Studien zu dieser Hypothese (Schellberg & Gembris, 2003 und Kopiez & Lehmann, 2008), die in Deutschland durchgeführt worden sind und die auf den ersten Blick unterschiedliche Ergebnisse zeigten. Die Daten, die diesen Studien zugrunde liegen, werden reanalysiert. Für jedes Paar von Klassengemeinschaften wird ein statistischer Homogenitätstest durchgeführt. Es kann gezeigt werden, dass sich die Präferenzverteilungen verschiedener Klassengemeinschaften öfter unterscheiden als unter der Annahme identischer Verteilungen zu erwarten wäre. Die statistischen Unterschiede zwischen den beiden Studien sind jedoch nicht größer als die Unterschiede zwischen den Klassengemeinschaften innerhalb beider Studien. Deshalb scheint die Klassengemeinschaft ein wichtiger Faktor für die Entwicklung musikalischer Präferenzen zu sein.

Abstract

The Hypothesis of open-earedness assumes that younger children are more open to unconventional music than older ones. This contribution is motivated by two studies on this hypothesis (Schellberg & Gembris, 2003 and Kopiez & Lehmann, 2008) that have been conducted in Germany and showed different results on first glance. The data these studies were based on is reanalyzed. For each pair of classroom communities a statistical test of homogeneity is performed. It can be shown that the preference-distributions of different classroom communities differ more often than expected under the assumption of identical distribution. The statistical differences between the two different studies, however, are not larger than the differences between the classroom communities within both of the studies. Therefore the classroom community seems to be an important factor in the development of musical preferences.

1 Einleitung

Seit den 80er Jahren wird vermutet, dass die Toleranz von Kindern gegenüber ihnen unbekannten Musikstilen im Laufe der Grundschulzeit abnimmt. Dieser vermutete Effekt wurde von Hargreaves als „Offenohrigkeit“ (open-earedness) bezeichnet (Hargreaves, 1982).

Schellberg und Gembris lieferten (2003) einen ersten empirischen Nachweis im deutschsprachigen Raum für einen Offenohrigkeitseffekt. Dieser Beitrag ist eine Reanalyse des Datenmaterials von zwei empirischen Untersuchungen zu den Musikpräferenzen von Kindern im Grundschulalter.

Die erste Untersuchung wurde von Schellberg und Gembris im Zeitraum von Mai bis Juli 2001 während des planmäßigen Unterrichts in einer Grundschule in Gaimersheim (Oberbayern) durchgeführt. Es nahmen 517 Grundschul Kinder an der Studie teil. Die Schüler hörten acht verschiedene Musikbeispiele von gut 1,5 Minuten Dauer (siehe Tabelle 1) und konnten auf einem Fragebogen angeben, wie gut ihnen die Musikbeispiele gefallen haben. Die hierfür verwendete 5-teilige Skala reichte von „will ich nicht hören“ bis „mag ich sehr gerne hören“ und wurde mit einer Smiley-Skala unterlegt, um auch von Kindern mit geringen Lesefertigkeiten verstanden zu werden. Der Datensatz dieser Untersuchung wird im Folgenden kurz mit SG 03 bezeichnet.

Die zweite Untersuchung wurde von Kopiez und Kobbenbring 2005 an zwei Grundschulen im Raum Hannover durchgeführt. An dieser Studie nahmen 188 Schüler teil. Das Versuchsdesign von Schellberg und Gembris wurde übernommen, um Vergleichbarkeit herzustellen (dabei wurde allerdings das Musikbeispiel für aktuelle Popmusik variiert, siehe Tabelle 1). Der Datensatz dieser Untersuchung wird im Folgenden mit KL08 bezeichnet.

Schellberg und Gembris (2003) konnten in ihren Daten (SG 03) einen hochsignifikanten Alterseffekt feststellen, indem sie die Mediane der Schülerurteile der verschiedenen Klassenstufen für jedes einzelne Musikbeispiel miteinander verglichen. Außerdem fanden sie hochsignifikante Geschlechterunterschiede.

Kopiez und Kobbenbring (2006) wiesen in ihrem Datenmaterial (KL08) mithilfe einer Varianzanalyse (Generalized Linear Model) u. a. ebenfalls einen hochsignifikanten Alterseffekt nach; allerdings konnten durch diesen nur 2 % der Varianz aufgeklärt werden. Der größte Teil der Varianz blieb unaufgeklärt. Sie schreiben:

„In Übereinstimmung mit früheren Studien wurde populäre Musik durchgängig am meisten bevorzugt. Es gab jedoch keine extremen Ablehnungen anderer weniger bekannter Musikstile. [...] Der eigentliche Effekt einer altersbedingten Offenohrigkeit gegenüber unbekannten Musikstilen bleibt damit in dieser Studie klein und liegt mit 2 % aufgeklärter Varianz auf der Höhe der Interaktion zwischen Geschlecht und Stil.“

In einer weiteren Veröffentlichung untersuchten Gembris und Schellberg (2007) ihr Datenmaterial (SG 03) erneut und stellten mithilfe des verteilungsfreien Kruskal Wallis H-Test hochsignifikante Alterseffekte bei allen acht Musikbeispielen für jeweils aufeinanderfolgende Altersstufen fest. Dabei wurden die Alterseffekte im Gegensatz zu Schellberg und Gembris (2003) nicht nach Klas-

senstufe, sondern nach Lebensalter analysiert. Außerdem konnten mithilfe des Chi-Quadrat-Tests signifikante Geschlechterunterschiede festgestellt werden.

Kopiez und Lehmann (2008; vgl. auch Lehmann & Kopiez, 2011) untersuchten die Datensätze KL08 und SG 03 ein weiteres Mal, indem sie die Musikbeispiele zu einer Gruppe von konventioneller Musik (alte und aktuelle Popmusik) sowie einer Gruppe von unkonventioneller Musik zusammenfassten und die Mittelwerte der entsprechenden Präferenzurteile bildeten. Mithilfe eines t-Tests wurden die Mittelwerte aufeinanderfolgender Klassenstufen verglichen. Es ergab sich nur zwischen der ersten und zweiten Klasse eine signifikante Änderung der Gefallensurteile. Die Erstklässler beurteilten die unkonventionelle Musik besser als die Zweitklässler. Kopiez und Lehmann variierten ihre Auswertung, indem sie die klassische Musik aus der Menge der unkonventionellen Musik ausschlossen. Dadurch waren in KL08 keine signifikanten Alterseffekte mehr nachweisbar.

Warum gelingt es Kopiez und Lehman (2008) nur dann einen Alterseffekt nachzuweisen, wenn sie die Präferenzurteile verschiedener Musikbeispiele mitteln und die Mozart-Arie zur unbekannten Musik hinzuzählen? Warum klärt die Klassenstufe in Kopiez und Kobbenbring (2006) nur einen geringen Anteil der Varianz auf? Liegt es daran, dass Schellberg und Gembris nicht-parametrische Verfahren verwenden, während Kopiez und Lehmann mit verteilungsgebundenen Tests arbeiten? Sind die Datensätze SG 03 und KL08 verschieden? Und falls ja, worin besteht ihre Verschiedenheit?

Um diese Fragen zu klären, wurden zunächst explorative Analysen des Datenmaterials mithilfe verschiedener Data-Mining-Verfahren durchgeführt. Dies lieferte allerdings keine verwertbaren Resultate: Entscheidungsbäume (Quinlan, 1993) zeigten nur wenige Elemente in den terminalen Knoten, aber dennoch viele Fehlklassifikationen. Clusterverfahren (Dempster et al., 1977) führten nur zu Clustern geringer Qualität. Hingegen fiel schon optisch ins Auge, dass verschiedene Klassengemeinschaften innerhalb eines Jahrganges dasselbe Musikbeispiel sehr unterschiedlich bewerteten.

Dieser zunächst nur rein optische Befund lässt sich gut damit begründen, dass die Klassengemeinschaft mit vielen Faktoren zusammenhängt, die musikalische Werturteile beeinflussen, wie etwa soziale Herkunft, Musiklehrer, Instrumentalunterricht oder allgemeine Wertvorstellungen. Deshalb wird in diesem Beitrag eine Analyse auf dem Niveau der Klassengemeinschaften vorgenommen. Dieses Vorgehen erlaubt einen differenzierteren Blick auf das Verschwinden der Offenohrigkeit in der Grundschulzeit.

2 Fragestellungen und Hypothesen

Die zentrale Hypothese dieses Beitrags lautet: Die einzelnen Klassengemeinschaften innerhalb eines Jahrganges unterscheiden sich in ihren Präferenzurteilen ganz erheblich. Aus dieser Hypothese leiten sich die folgenden Fragestellungen ab:

- 1) Sind die Präferenzurteile bezüglich der verschiedenen Musikbeispiele in den einzelnen Klassengemeinschaften eines Jahrganges identisch verteilt oder lässt sich diese Hypothese signifikant ablehnen? (siehe Abschnitt 4.1)

- 2) Gibt es Teilmengen von Klassengemeinschaften, die bezüglich ihrer Präferenzurteile einander besonders ähnlich oder besonders unähnlich sind? (siehe Abschnitt 4.2)
- 3) Spielt es für die Ähnlichkeit der Präferenzurteile zweier Klassengemeinschaften eine Rolle, ob sie aus dem gleichen Datensatz oder aus unterschiedlichen Datensätzen stammen? (siehe Abschnitt 4.2)

Im Verlaufe des Reanalyseprozesses ergaben sich die folgenden weiteren Fragestellungen:

- 4) Sind Musikbeispiele, die aus musikwissenschaftlicher Sicht unterschiedlichen Genres zuzuordnen sind, für die Kinder auch verschieden, oder unterscheiden sich die kindlichen Begriffe musikalischer Ähnlichkeit von den üblichen Genrezuordnungen? (siehe Abschnitt 4.3 und 4.4)
- 5) Wie homogen sind die Präferenzurteile innerhalb einer Klassengemeinschaft? (siehe Abschnitte 4.5 und 4.6)

3 Methode

Da Kopiez und Lehmann die Versuchsanordnung und das Stimulusmaterial von Schellberg und Gembris übernommen haben, war es möglich das Datenmaterial der beiden Untersuchungen zusammenzuführen und gemeinsam zu analysieren. Bei den Musikbeispielen zur aktuellen Popmusik ist allerdings zu beachten, dass in den beiden Studien aufgrund der unterschiedlichen Erhebungszeitpunkte unterschiedliches Stimulusmaterial verwendet wurde (siehe Tabelle 1).

Tab. 1:
Liste der Musikbeispiele

Komponist/Interpret	Titel	Genre
Bulgarian Voices Angelite	Dancing Voices (trad.)	Ethnische Musik
Bach	Suite Nr. 3, D-Dur, Gavotte 1	Klassik
Mozart	Arie: „Voi, che sapete“ aus der „Hochzeit des Figaro“	Klassik
Mendelssohn-Bartholdy	Sinfonie Nr. 4 op. 90, 1. Satz	Klassik
Propaganda (1990)	„Heaven Give Me Words“	Popmusik (nicht aktuell)
Scelsi	Canti del Capricorno, Nr. 1	Avantgarde
Henze	3. Sinfonie, 3. Satz Beschwörungstanz	Avantgarde
Marque (2000)	„One To Make Her Happy“	Popmusik (aktuell SG 03)
Daniel Powter (2005)	„Bad Day“	Popmusik (aktuell KL08)

Jedes Musikbeispiel dauert 1 bis 1,5 Minuten, die Schüler können auf einer fünfteiligen Skala angeben, wie gut ihnen die Musik gefällt. Die Werturteile sind auf einer 5-teiligen Skala ordinal skaliert von „mag ich sehr gerne hören“ (1) bis „will ich nicht hören“ (5). Schellberg und Gembris haben im Gegensatz zu Kopiez und Lehmann neben Grundschulern auch Daten von Kindergartenkindern und den Jahrgangsstufen 5 und 6 erhoben. Aus Gründen der Vergleichbarkeit wurden in der vorliegenden Untersuchung aber nur die Daten der Klassenstufen 1–4 betrachtet. Das Datenmaterial umfasst fünf Klassenzüge (bezeichnet mit a bis e) aus der Studie von Schellberg und Gembris, die 2000 in Gaimersheim (Oberbayern) durchgeführt wurde und zwei Klassenzüge (bezeichnet mit A und B) aus der Erhebung von Kopiez und Lehmann 2005 in Hannover, also insgesamt sieben Klassengemeinschaften in jedem Jahrgang. Diese teilen sich auf in 182 Schüler aus der ersten Jahrgangsstufe, 179 Schüler aus der zweiten Jahrgangsstufe, 171 Schüler aus der dritten Jahrgangsstufe und 173 Schüler aus Jahrgangsstufe vier. Das sind insgesamt 705 Schüler¹.

In Tabelle 2 sind die Größen der Klassengemeinschaften verzeichnet. Sie bilden die Stichprobenumfänge der nachfolgend beschriebenen Homogenitätstests.

Tab. 2:
Matrix der Klassengrößen in den Datensätzen (SG 03 und KL08)

	a	b	c	d	e	A	B	Summe
1	29	28	25	24	25	22	29	182
2	24	29	27	26	26	23	24	179
3	26	23	26	27	27	19	23	171
4	27	25	22	26	25	21	27	173
Summe	106	105	100	103	103	85	103	705

Anmerkung: Die Stichprobenumfänge in den Homogenitätstests sind gleich der Anzahl der Schüler in den betreffenden Klassengemeinschaften. In den Spalten sind die jeweiligen Klassenzüge verzeichnet, in den Zeilen die Jahrgangsstufen.

Die Auswertung der Daten aus SG 03 und KL08 auf dem Niveau der Klassengemeinschaften unterliegt einigen Einschränkungen:

- 1) Die Daten sind ordinal skaliert. Daher sind verschiedene statistische Maßzahlen wie Mittelwert, Varianz und Korrelation nicht anwendbar. Andere Lokations- und Streuungsmaße wie Median und Quantilsabstand haben jedoch den Nachteil, nur wenige verschiedene Werte – nämlich die Zahlen von eins bis fünf – annehmen zu können und führen daher zu einem großen Informations-

¹ Die Anzahl der VPn aus KL08 summiert sich zu insgesamt 188. Das sind 2 VPn mehr als von Kopiez & Kobbenbring (2006) bzw. Kopiez & Lemann (2008) angegeben. Dort sind diese beiden VPn aus Klasse 2A wegen missing values bei den Musikbeispielen von Henze und Propaganda ausgeschlossen worden. Dies war hier nicht nötig. Hier blieben nur die beiden missing values unberücksichtigt.

verlust. Für die Beurteilung der Urteilshomogenität innerhalb einer Klassengemeinschaft wird daher die Shannon-Information (Shannon & Weaver, 1949) verwendet. Die Shannon-Information ist ein Streuungsmaß für nominal oder ordinal skalierte Merkmale (vgl. Rinne, 2003, S. 46). Sie quantifiziert die Ähnlichkeit einer diskreten Wahrscheinlichkeitsverteilung auf n Punkten mit der Gleichverteilung und nimmt Werte im Intervall zwischen 0 und $\log_2 n$ an. Der minimale Wert 0 besagt, dass die Verteilung auf einem einzigen Punkt konzentriert ist (Dirac-Verteilung). Der maximale Wert $\log_2 n$ zeigt an, dass alle Punkte gleichwahrscheinlich sind (Gleichverteilung)². Da die Gefallensurteile in SG 03 und KL08 fünf verschiedene Ausprägungen annehmen können, liegt die Information einer Präferenzverteilung immer im Intervall zwischen 0 und $\log_2 5 \approx 2,32$.

- 2) Bei einer Analyse auf dem Niveau der Klassengemeinschaften sind die Stichprobenumfänge jeweils gleich der Anzahl der Schüler in einer Klasse; sie liegen also zwischen 19 und 29 (vgl. Tabelle 2). Bei solch kleinen Stichprobenumfängen sind statistische Verfahren, die auf den zentralen Grenzwertsatz (vgl. etwa Gänssler & Stute, 1977, S. 154) zurückgehen (wie z. B. der Chi-Quadrat-Test) nicht anwendbar ohne zuvor einen bedeutenden Informationsverlust durch Datenaggregation hinzunehmen. Deshalb wird hier auf die Kolmogorov-Smirnov-Statistik zurückgegriffen, die auf dem elementarerem Satz von Glivenko und Cantelli (vgl. etwa Gänssler & Stute, 1977, S. 145), nämlich der Konvergenz der relativen Häufigkeiten gegen die theoretischen Wahrscheinlichkeiten, beruht.

Zunächst wurden die Präferenzverteilungen der sieben Klassengemeinschaften für jedes Musikbeispiel und für jeden Jahrgang einzeln paarweise verglichen. Es ergaben sich so $21 = 7(7-1)/2$ Paarungen von Klassengemeinschaften³ für jedes der acht Musikbeispiele und für jeden der vier Jahrgänge. Für jede Paarung wurde ein zweiseitiger Kolmogorov-Smirnov-Homogenitätstest auf dem Niveau $p > 0,90$ durchgeführt (vgl. Rinne, 2003, S. 563 f.). Das Niveau $p > 0,90$ wurde gewählt, um Deckeneffekte (nur Ablehnungen oder nur Annahmen) zu vermeiden.

Hierzu wurde die Nullhypothese aufgestellt, dass die zwei Klassengemeinschaften i und j sich in ihrer Präferenzverteilung hinsichtlich eines Musikbeispiels nicht unterscheiden:

$$F_i(x) = F_j(x).$$

Diese Nullhypothese wurde gegen die Alternative $F_i(x) \neq F_j(x)$ getestet, indem die Prüfgröße

$$D(n_i, n_j) = \max_x \{ |F_i(x) - F_j(x)| \}$$

2 Die Verwendung der Shannon-Information hat rein statistische Gründe und steht in keinerlei Beziehung zur informationstheoretischen Ästhetik der 70er Jahre.

3 Will man n Punkte durch genau eine Linie miteinander verbinden, so benötigt man hierzu $n(n-1)/2$ Verbindungslinien. Dies lässt sich mit vollständiger Induktion leicht beweisen. Auf die gleiche Anzahl kommt man, wenn man die Elemente unterhalb (bzw. oberhalb) der Diagonalen einer $n \times n$ -Matrix durchzählt.

berechnet wurde, und diese auf die Ablehnungsbedingung

$$D(n_i, n_j) > 1,22 \times \left(\sqrt{\frac{n_i + n_j}{n_i \times n_j}} \right)$$

überprüft wurde (vgl. Massey, 1952), wobei n_i und n_j die in Tabelle 2 verzeichneten Stichprobenumfänge (bzw. Klassengrößen) bezeichnen und F_i bzw. F_j die empirischen Verteilungsfunktionen der Klassengemeinschaften i und j (siehe Abb. 1). Führt dieser Test zur Ablehnung, so kann man daraus schließen, dass den Präferenzverteilungen der beiden betreffenden Klassengemeinschaften mit einer Wahrscheinlichkeit von $p > 0,9$ nicht das gleiche stochastische Phänomen zugrunde liegt; die Wahrscheinlichkeit dafür, dass die Präferenzverteilungen identisch sind, ist hingegen kleiner als 0,1.

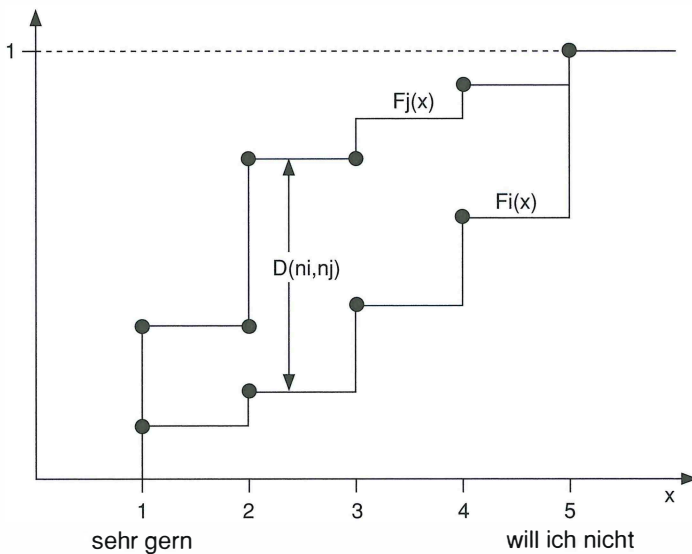


Abb. 1:

Zwei empirische Verteilungsfunktionen und die Kolmogorov-Smirnov-Statistik. Eine empirische Verteilungsfunktion nimmt an der Stelle x den Wert der kumulierten relativen Häufigkeiten, die kleiner als x sind, an.

Man erhält so für jede Kombination aus Musikbeispiel und Jahrgangsstufe eine Matrix, in der verzeichnet ist, für welche Paarungen von Klassengemeinschaften der Homogenitätstest zur Ablehnung oder Annahme der Nullhypothese $F_i(x) = F_j(x)$ führt. Diese Matrix geht auf die Durchführung von $7 \times 7 = 49$ Homogenitätstests zurück.

Um die Ähnlichkeit der einzelnen Klassengemeinschaften bezüglich ihrer Präferenzurteile besser beurteilen zu können, wird für jedes Paar von Klassengemeinschaften die Anzahl der Musikbeispiele gezählt, für die der Homogeni-

tätstest nicht abgelehnt wird. Da es 7 Klassenzüge und 8 Musikbeispiele gibt, entsteht hierdurch für jede Jahrgangsstufe eine symmetrische 7×7 -Matrix mit ganzzahligen Einträgen von 0 bis 8. Diese Matrizen gehen jeweils auf die Durchführung von $8 \times 49 = 392$ Homogenitätstests zurück und werden im Folgenden als Klassen-Präferenzmatrix bezeichnet. Die Homogenitätstests dienen also hier nicht dazu, eine bestimmte Hypothese empirisch zu untermauern, sondern um Ähnlichkeit von Präferenzurteilen zu quantifizieren. Dieses Ähnlichkeitsmaß ist differenzierter als z. B. der Median, abstrahiert aber im Gegensatz zu diesem von dem mittleren Wert der Präferenz, ob also ein Musikbeispiel einer Klassengemeinschaft mehr oder weniger gefällt.

Zur besseren zweidimensionalen Darstellung der Ähnlichkeitsverhältnisse wird eine Hauptachsentransformation (siehe etwa Bronstein et al., 2001, S. 292) durchgeführt. Mithilfe der Hauptachsentransformation, auch Eigenwertzerlegung (eigenvalue decomposition) genannt, wird eine symmetrische Matrix als Produkt von einer Diagonalmatrix L und zwei orthonormalen Matrizen B und B^T dargestellt:

$$M = BLB^T$$

Die Diagonalmatrix L enthält auf der Diagonalen die Eigenwerte der Matrix M . Sie bewirkt also eine Streckung bzw. Stauchung des Raumes in Richtung der Einheitsvektoren. Die orthonormalen Matrizen, die aus normierten Eigenvektoren von M zusammengesetzt sind, bewirken Drehungen und Spiegelungen des Koordinatensystems.

Dieses Verfahren ähnelt der Faktorenanalyse, mit dem Unterschied, dass die Hauptachsentransformation die Verzerrungen der Abstände zwischen den Datenpunkten, die durch die Dimensionsreduktion unvermeidlich ist, minimiert (Eckart & Young, 1936). Die Faktorenanalyse wird hingegen üblicherweise auf Korrelationsmatrizen angewendet und zielt darauf ab, die Faktoren so zu wählen, dass sie möglichst viel Varianz aufklären (vgl. etwa Tränkle, 1991, S. 250). Da es sich bei den Einträgen in den Matrizen um Abstände bzw. Ähnlichkeiten handelt und nicht um Korrelationen oder Kovarianzen, ist die Hauptachsentransformation angemessen.

Die zweidimensionale Darstellung der Ähnlichkeitsverhältnisse erhält man dadurch, dass jede Klassengemeinschaft durch die entsprechenden Komponenten der Eigenvektoren zu den beiden größten Eigenwerten repräsentiert wird. Dies entspricht bei der Faktorenanalyse der zweidimensionalen Darstellung der Faktoren mit der höchsten Faktorladung.

Ebenso wie für Paare von Klassengemeinschaften wird für Paare von Musikbeispielen verfahren. Auf diese Weise wird Aufschluss über die Ähnlichkeit der Musikbeispiele bezüglich ihrer Präferenzurteile in den einzelnen Klassengemeinschaften gewonnen. Man erhält für jede Jahrgangsstufe eine symmetrische 8×8 -Matrix mit ganzzahligen Einträgen von 0 bis 7, die jeweils angeben, wie viele Klassengemeinschaften ein Paar von Musikbeispielen signifikant unterschiedlich bewerten. Diese Matrizen werden im Folgenden als Musik-Präferenzmatrix bezeichnet. Auch hier wird eine Hauptachsentransformation

durchgeführt, um die Ähnlichkeitsverhältnisse zweidimensional darstellen zu können.

Alle Berechnungen für diesen Beitrag wurden mit der Open-Source-Software R durchgeführt⁴.

4 Ergebnisse

4.1 Unterschiedliche Präferenzverteilungen zwischen den Klassengemeinschaften

Um zu überprüfen, ob die Klassengemeinschaften innerhalb eines Jahrgangs sich in ihren Präferenzurteilen für ein und dasselbe Musikbeispiel unterscheiden, wurde für jedes der 21 Paare von Klassengemeinschaften die Summe der abzulehnenden Homogenitätstests gebildet. Die Ergebnisse finden sich in Tabelle 3.

Bei der Beurteilung der Werte in Tabelle 3 muss man sich vergegenwärtigen, dass die Ablehnung eines statistischen Tests immer nur eine Wahrscheinlichkeitsaussage ist. Wird das Signifikanzniveau $p > 0,90$ gewählt, so bedeutet dies, dass eine Ablehnung der Nullhypothese nur mit einer Wahrscheinlichkeit von 90 % korrekterweise erfolgt. Es bleibt eine Wahrscheinlichkeit von 10 % dafür, dass die Teststatistik die Ablehnungsbedingung erfüllt, obwohl die Nullhypothese richtig gewesen wäre. Anders gesprochen bedeutet dies: Die irrtümliche Ablehnung ist Bernoulli-verteilt zum Parameter $q = 0,1$.

Führt man nun unabhängig voneinander mehrere Homogenitätstests durch, so ist die Summe der irrtümlichen Ablehnungen so verteilt wie die Summe der entsprechenden Bernoulli-Variablen. Die Verteilung dieser Summe wird durch eine Binomialverteilung⁵ beschrieben. Genauer: Eine Summe von n unabhängigen Bernoulli-Variablen mit dem Parameter q ist binomial-verteilt mit den Parametern n und q .

Da in Tabelle 3 jeweils verzeichnet ist, für wie viele der 21 Paare von Klassengemeinschaften der Homogenitätstest auf dem Niveau $p > 0,90$ abgelehnt wird, ist die Anzahl der irrtümlichen Ablehnungen in jedem Eintrag in Tabelle 3 binomial-verteilt mit den Parametern $n = 21$ und $q = 0,1$.

Daraus lässt sich berechnen, dass ein Tabellenwert ≥ 5 unter der Annahme von H_0 (alle Präferenzurteile der Klassengemeinschaften eines Jahrgangs seien für ein Musikbeispiel identisch verteilt) nur mit einer Wahrscheinlichkeit von weniger als 5 % auftreten kann, ein Tabellenwert von ≥ 6 nur mit einer Wahrscheinlichkeit von weniger als 1,5 % und ein Tabellenwert von ≥ 7 nur mit einer Wahrscheinlichkeit von 0,5 %. Tabellenwerte ≥ 5 indizieren also signifikante Unterschiede $p > 0,95$ in den Präferenzverteilungen und Tabellenwerte ≥ 7 hochsignifikante Unterschiede $p > 0,995$.

4 R Development Core Team (2009). R: A Language and Environment for Statistical Computing: Wien. <http://www.R-project.org> (Frühjahr 2011).

5 Zur Definition der Binomialverteilung siehe z. B. Rinne (2003, S. 253).

Die für die Paare von Klassengemeinschaften durchgeführten Homogenitätstests zeigen, dass den Präferenzverteilungen der einzelnen Klassengemeinschaften innerhalb eines Jahrganges tatsächlich unterschiedliche stochastische Phänomene zugrunde liegen. Daraus kann man schließen, dass die Klassengemeinschaften selber einen Einfluss auf die Präferenzurteile haben⁶.

Auffällig ist der hohe Wert für Mozart in der ersten Klasse, die völlige Homogenität der Popmusik in den Klassen 2 und 3, sowie die starke Zunahme der Homogenität für den bulgarischen Chor in der vierten Klasse. Alle anderen Musikbeispiele weisen eine leichte Zunahme der Unterschiede in den Präferenzverteilungen auf. Interessant ist auch, dass sich die Klassengemeinschaften des ersten Jahrgangs in ihren Urteilen zur Avantgardemusik (Musikbeispiele von Scelsi und Henze) nicht unterscheiden. Erst in den folgenden Jahrgängen nehmen die Klassengemeinschaften für die Avantgardemusik unterschiedliche Präferenzverteilungen an.

Die Homogenität der Präferenzurteile für die aktuelle Popmusik in den Klassen 2 und 3 ist umso bemerkenswerter, als sich die Präferenzurteile in den beiden zugrundeliegenden Datenerhebungen auf unterschiedliches Stimulusmaterial beziehen, das zu den unterschiedlichen Erhebungszeitpunkten jeweils in den aktuellen Charts war.

Tab. 3:
Anzahl abgelehnter Homogenitätstests in Abhängigkeit von Jahrgang und Musikbeispiel

Jahrgang	Chor	PopOld	Mendel.	Scelsi	Mozart	Bach	Henze	PopNew
1. Klasse	4	5	4	1	13	5	2	4
2. Klasse	6	5	7	6	3	9	5	0
3. Klasse	11	7	8	8	5	8	6	0
4. Klasse	1	7	10	6	4	9	9	10

Anmerkung: Hier ist dargestellt, für wie viele der 21 Klassenpaarungen die Homogenitätsannahme auf dem Niveau $p < 0,90$ abgelehnt wird. Bei identischer Verteilung sind 2,1 Ablehnungen zu erwarten.

Die starke Zunahme der Homogenität für den bulgarischen Chor von der dritten zur vierten Klassenstufe ist vermutlich ein Nebeneffekt der Uneinheitlichkeit der Präferenzurteile innerhalb der einzelnen Klassengemeinschaften. In der vierten Klasse nimmt die Information der Präferenzverteilungen für den bulgarischen Chor einen Wert von über 2,2 Bit an. Die Verteilungen ähneln daher der Gleichverteilung (mit einem Wert von $\log_2 n \approx 2,32$) und sind daher in fast allen Klassengemeinschaften ähnlich.

6 Diese Aussage würde sich im Prinzip nicht ändern, wenn für die Kolmogorov-Smirnov-Tests ein anderes Signifikanzniveau gewählt worden wäre. Dann wären die Werte in Tabelle 3 anders, aber die Parameter der Binomialverteilung würden sich entsprechend ändern.

Die große Unterschiedlichkeit der Präferenzverteilungen für das Musikbeispiel von Mozart in der ersten Klasse geht tatsächlich auf einen signifikanten Unterschied zwischen den Daten von SG 03 und KL08 zurück. Während die Belcanto-Arie von Mozart in KL08 vorwiegend positiv beurteilt wird, erhält sie in SG 03 vorwiegend negative oder ausgeglichene Beurteilungen (siehe Tabelle 4). Dies könnte ein Grund dafür sein, dass Kopiez und Lehmann (2008) nur dann einen signifikanten Offenohrigkeitseffekt beobachten konnten, wenn sie die klassische Musik (also auch das Musikbeispiel von Mozart) zur „ungewöhnlichen“ Musik hinzuzählten.

Tab. 4:
Präferenzverteilungen für das Musikbeispiel von Mozart in der ersten Klasse

	1a	1b	1c	1d	1e	1A	1B
++	9	8	4	1	7	16	14
+	1	2	4	4	10	1	2
0	8	4	3	2	3	1	9
–	6	1	3	2	2	2	2
--	5	13	11	15	3	2	2

Anmerkung: Es zeigt sich eine Gruppe von Mozartliebhabern (1e, 1A und 1B) und eine Gruppe von solchen Klassengemeinschaften, die Mozart ablehnen (1b, 1c und 1d).

4.2 Geringe Unterschiede zwischen den Präferenzurteilen in SG 03 und KL08

Anhand der Klassenpräferenzmatrizen kann man feststellen, dass die Unterschiede zwischen den beiden Studien nicht größer sind als zwischen den Klassengemeinschaften innerhalb der einzelnen Studien. In den ersten drei Jahrgangs-

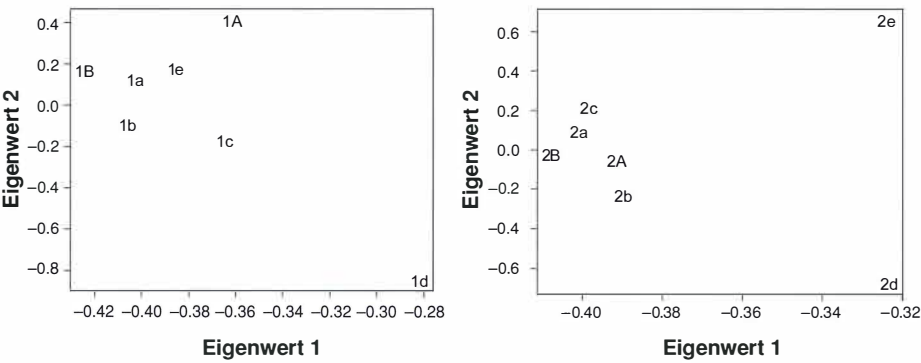


Abb. 2:
Hauptachsentransformation der Klassen-Präferenzmatrizen in der ersten und zweiten Klassenstufe. Es zeigt sich jeweils ein Hauptcluster und eine bzw. zwei Ausreißerklassen, die jeweils aus SG 03 stammen.

stufen zeigen sich jeweils Ausreißerklassen (1d, 2d, 2e und 3c), diese stammen durchweg aus einer der fünf Klassengemeinschaften von SG 03 (siehe Abbildung 2 und 3). Auch in der vierten Jahrgangsstufe liegen die Klassengemeinschaften aus KL08 jeweils näher an einer Klassengemeinschaft aus SG 03 als an der jeweils anderen Klasse aus KL08. Die Unterschiede in den Ergebnissen der Studien von Schellberg und Gembris einerseits sowie Kopiez und Kobbenbring bzw. Kopiez und Lehman andererseits lassen sich also dadurch erklären, dass sie auf den Präferenzurteilen unterschiedlicher Klassengemeinschaften beruhen. Klassengemeinschaften mit einem eindeutigen Offenohrigkeitseffekt gibt es zufälligerweise nur in SG 03. In der ersten Klassenstufe von KL08 gibt es hingegen viele ‚Mozartliebhaber‘.

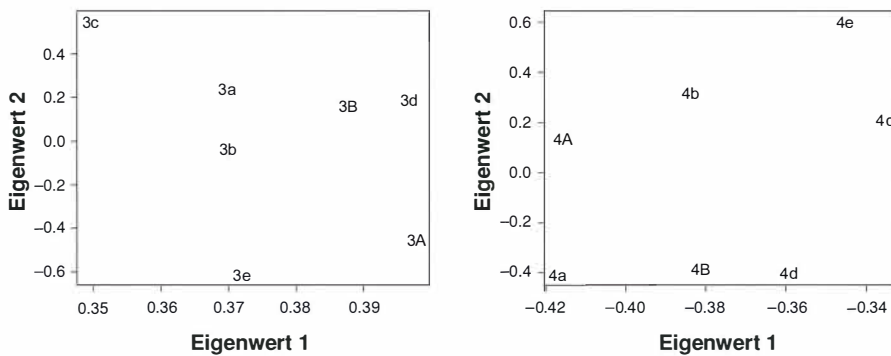


Abb. 3:

Hauptachsentransformation der Klassen-Präferenzmatrizen in der dritten und vierten Klassenstufe. Die Klassengemeinschaften aus KL08 unterscheiden sich nicht grundsätzlich von denen aus SG 03.

4.3 Zunehmende Differenziertheit der Präferenzurteile

Betrachtet man die Präferenzverteilungen für Paare von Musikbeispielen aggregiert über die Klassengemeinschaften, so stellt man fest, dass die Anzahl der Paarungen von Musikbeispielen, für die man den Homogenitätstest ablehnen kann, von der ersten bis zur vierten Klasse zunimmt, und zwar ganz besonders von der ersten zur zweiten Klasse. Erhalten in der ersten Klasse nur 32 % der Paare von Musikbeispielen unterschiedliche Präferenzverteilungen, so sind dies in der zweiten Klasse schon 56 %, in der dritten Klasse 57 % und in der vierten Klasse schließlich 64 %. Die Erstklässler machen in ihren Präferenzurteilen innerhalb einer Klassengemeinschaft also nur einen sehr geringen Unterschied zwischen den verschiedenen Musikbeispielen. Erst in den folgenden Klassen werden die verschiedenen Musikbeispiele innerhalb einer Klassengemeinschaft unterschiedlich beurteilt.

Gembris und Schellberg (2007) ziehen in ihrer Studie das Fazit: „Grob vereinfachend könnte man sagen: je jünger die Kinder, desto besser kommt Musik

bei ihnen an, im Prinzip jedenfalls.“ Die hier gewonnenen Ergebnisse legen eine andere Sichtweise nahe: Je jünger die Kinder, desto geringer die Unterschiede in der Bewertung verschiedener Musikbeispiele.

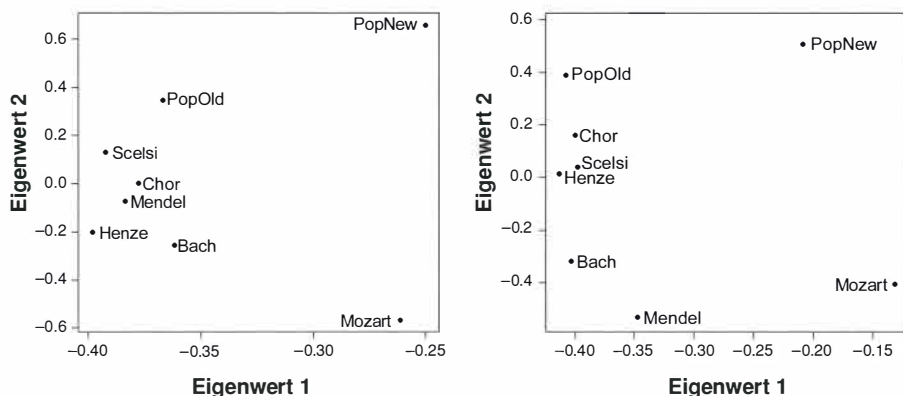


Abb. 4:

Hauptachsentransformation der Musik-Präferenzmatrizen in der ersten und zweiten Klassenstufe. In der ersten Klasse (links) finden sich alle Musikbeispiele außer Mozart und der aktuellen Popmusik in einem Cluster. In der zweiten Klasse (rechts) erkennt man, dass sich hiervon ein Cluster klassischer Instrumentalmusik abspaltet.

4.4 Entwicklung musikalischer Ähnlichkeitsbegriffe

Ein Blick auf die Hauptachsentransformation der Musik-Präferenzmatrix in Abbildung 4 ergibt ein genaueres Bild der Entwicklung musikalischer Ähnlichkeitsbegriffe. In der ersten Klasse erscheinen Mozart und die aktuelle Popmusik von den anderen Musikbeispielen getrennt. Alle anderen Musikbeispiele finden sich in einem Cluster. Es scheint, als hätten die Kinder in diesem Alter zwar einen Begriff von aktueller Popmusik und Belcanto, aber noch keine weiter differenzierten Begriffe von Musik. So erhalten die Musikbeispiele von Mendelssohn und Scelsi in keiner Klassengemeinschaft unterschiedliche Präferenzurteile; Gleiches gilt für Henze und Bach: Die entsprechenden Einträge in der Musik-Präferenzmatrix enthalten jeweils den Wert 7. Schon in der ersten Klasse zeigt sich allerdings ein Befund, der in den weiteren Jahrgängen erhalten bleibt: Der Belcanto-Gesang von Mozart ist ein Gegenspieler der aktuellen Popmusik. Nur eine Klassengemeinschaft der ersten Jahrgangsstufe unterscheidet diese Musikbeispiele nicht signifikant.

In der zweiten Klasse spaltet sich ein Cluster mit klassischer Instrumentalmusik ab. Interessanterweise findet sich die Avantgardemusik mit der nicht-aktuellen Popmusik in einem Cluster. Dazu passen Kommentare der Versuchspersonen: „Bulgarisch Cool“ (Klasse 2A) oder „Scelsi zu einfalllos“ (Klasse 2A). Es scheint sich ein Begriff für klassische Musik herauszubilden. Zwischen nicht-aktueller Popmusik, Avantgarde und ethnischer Musik wird jedoch noch

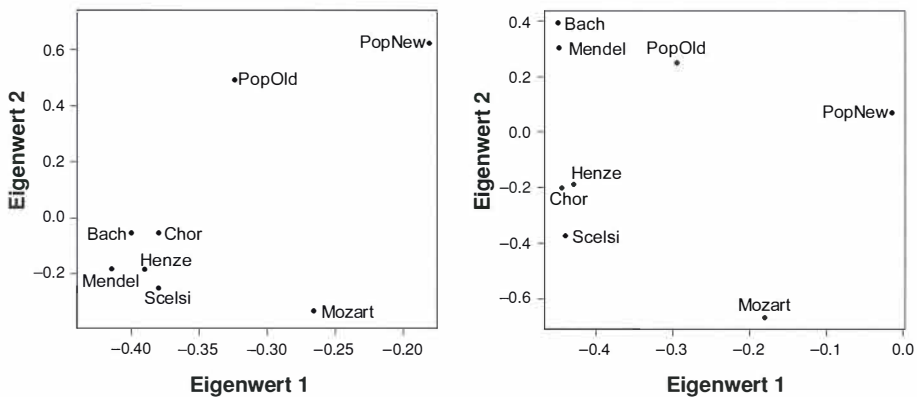


Abb. 5:

Hauptachsentransformation der Musik-Präferenzmatrizen für die Klassenstufen drei und vier. In der dritten Klasse (links) trennt sich die nicht-aktuelle Popmusik von den anderen Musikbeispielen. In der vierten Klasse (rechts) trennen sich schließlich auch Avantgarde und ethnische Musik einerseits von der klassischen Instrumentalmusik andererseits.

nicht differenziert. Die dritte Klassenstufe unterscheidet hingegen nicht-aktuelle Popmusik einerseits sowie Avantgarde, Klassik und ethnische Musik andererseits und fasst Letztere in einem Cluster zusammen. In der vierten Klasse schließlich clustern die Präferenzurteile so, wie es ihrer Genrezuordnung entspricht: Klassik, Avantgarde/Ethnic, Belcanto, unbekannte Popmusik und aktuelle Popmusik.

4.5 Unterschiedliche Urteilsfreiheit innerhalb der Klassengemeinschaften

Es wurde vermutet, dass die Klassengemeinschaften bezüglich unterschiedlicher Musikstile einen mehr oder weniger starken Einfluss auf die Präferenzurteile ihrer Individuen ausüben, dass also die Individuen innerhalb einer Klassengemeinschaft mehr oder weniger frei in ihren Präferenzurteilen sind. Diese Urteilsfreiheit wird durch die Shannon-Information der Präferenzverteilungen der einzelnen Klassengemeinschaften gemessen.

Wird über alle Musikbeispiele gemittelt, so zeigt sich, dass sich die Klassengemeinschaften in ihrer mittleren Shannon-Information stark unterscheiden. Es gibt Klassengemeinschaften mit sehr einheitlichen Präferenzurteilen (geringe Information) und solche mit sehr uneinheitlichen (große Information) (vgl. Abb. 5). Dies weist daraufhin, dass die Urteilsfreiheit in den einzelnen Klassen sehr unterschiedlich ist. Bildet man den Mittelwert über die einzelnen Klassen, so fällt auf, dass die Shannon-Information von der ersten zur zweiten Klasse deutlich zunimmt und danach nahezu konstant bleibt.

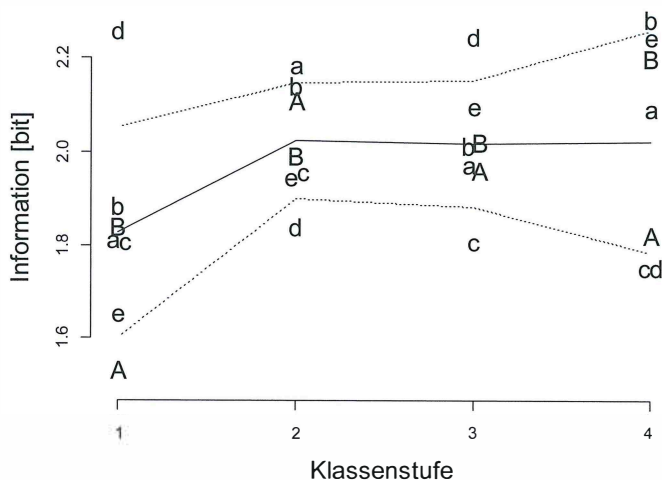


Abb. 6:

Shannon-Information der einzelnen Klassengemeinschaften gemittelt über alle Musikbeispiele. Großbuchstaben bezeichnen die Klassengemeinschaften aus KL08, Kleinbuchstaben die aus SG 03. Die durchgezogene Linie gibt die jeweiligen Mittelwerte an, die gepunktete die Standardabweichung.

4.6 Urteilsfreiheit bezüglich aktueller Popmusik

Die Shannon-Information für aktuelle Popmusik ist deutlich geringer als die für alle anderen Musikbeispiele, sie beträgt meist deutlich weniger als 1 Bit im Gegensatz zu den anderen Musikbeispielen, die Informationswerte von 1,4 bis 2,2 Bit aufweisen. Dies kann als Deckeneffekt interpretiert werden, weil die aktuelle Popmusik stets äußerst positive Urteile erhält, oder aber auch als Hinweis auf die soziale Gebundenheit der Präferenzurteile für aktuelle Popmusik. Es gibt allerdings einige Klassengemeinschaften (1b und 1d sowie 4e und 4b aus SG 03) in denen die Information für aktuelle Popmusik ungewöhnlich hoch ist. In diesen Klassen beurteilt ein großer Teil der Schüler die aktuelle Popmusik negativ oder neutral.

5 Diskussion

In diesem Beitrag konnte gezeigt werden, dass zwischen den Musikpräferenzen der verschiedenen Klassengemeinschaften innerhalb eines Jahrganges große Unterschiede bestehen, und dass es daher sinnvoll ist, die Klassengemeinschaft als unabhängige Variable zu betrachten.

Die Unterschiede zwischen den Datensätzen KL08 und SG 03 sind kleiner als die Variationsbreite innerhalb von SG 03. Die Tatsache, dass sich einige

Resultate, die anhand von SG 03 gewonnen wurden, anhand von KL08 nicht replizieren ließen, ist darauf zurückzuführen, dass unterschiedliche Klassengemeinschaften sehr unterschiedliche Präferenzverteilungen zeigen. Die Annahme einer weiteren Erklärungsvariable ist hierzu nicht notwendig. Es wäre daher wünschenswert, wenn mehr Information über die Klassengemeinschaften verfügbar wäre (etwa die überwiegende soziale Herkunft der Schüler, der Musiklehrer etc.), um die Unterschiede der Präferenzurteile erklären zu können.

In der ersten Klasse bilden sich zunächst Begriffe von aktueller Popmusik und Belcanto heraus. Alle anderen Musikbeispiele werden ähnlich beurteilt. Keine der 7 Klassen beurteilte Mendelssohn und Scelsi sowie Bach und Henze verschieden.

In der zweiten Klasse bildet sich ein Cluster mit „interessanter“ Musik heraus, in der ethnische und Avantgarde Musik ebenso vertreten ist wie die nicht-aktuelle Popmusik.

In der vierten Klasse schließlich clustern die Präferenzurteile so, wie es ihrer Genrezuordnung entspricht: Klassik, Avantgarde/Ethnic, Belcanto, unbekannte Popmusik und aktuelle Popmusik.

Diese Befunde stehen im Einklang mit einem Ergebnis von Gembris und Schellberg (2007), nach dem das Alter zwischen 8 und 9 Jahren eine kritische Phase für den kindlichen Musikgeschmack zu sein scheint. Hierzu passt auch die zunehmende Differenziertheit der Präferenzurteile von der ersten zur zweiten Jahrgangsstufe.

Zu dem diesem Beitrag zugrundeliegenden Datenmaterial ist kritisch anzumerken, dass es zu jedem Musikbeispiel nur ein einziges Item pro Versuchsperson, nämlich ein Präferenzurteil, enthält. Die gemachten Aussagen beruhen daher auf einer unsicheren empirischen Basis. Wünschenswert wäre ein Versuchsdesign, das Musikpräferenzen auf mehrere Items zurückführen lässt. Außerdem ist es fraglich, ob es gerechtfertigt ist, aus einem Präferenzurteil zu einem oder zwei Musikbeispielen eine Präferenz für ein musikalisches Genre herauszulesen. Auch in dieser Hinsicht wäre eine Weiterentwicklung des Versuchsdesigns erstrebenswert.

Literatur

- Bronstein, I. N., Semendjajew, K. A., Musiol, G. & Mühlig, H. (2001). *Taschenbuch der Mathematik*. Frankfurt a. M.: Harri Deutsch.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B*, 39 (1), 1–38.
- Eckart, C. & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1 (3), 211–218.
- Gänssler, P. & Stute, W. (1977). *Wahrscheinlichkeitstheorie*. Berlin, Heidelberg, New York: Springer.
- Gembris, H. & Schellberg, G. (2007). Die Offenohrigkeit und ihr Verschwinden bei Kindern im Grundschulalter. In W. Auhagen, C. Bullerjahn & H. Höge (Hrsg.), *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie* (Bd. 19, S. 71–92). Göttingen: Hogrefe.

- Hargreaves, D.J. (1982). The development of aesthetic reactions to music. *Psychology of Music (Special Issue)*, 51–54.
- Kopiez, R. & Kobbenbring, M. (2006). Die Altersabhängigkeit des Musikgeschmacks: neue Ergebnisse zur Hypothese der „Offenohrigkeit“. *Tagungsband der DGM-Jahrestagung 2006*, 39–41.
- Kopiez, R. & Lehmann, M. (2008). The ‚open-earedness‘ hypothesis and the development of age-related aesthetic reactions to music in elementary school children. *British J. Music Education* 25 (2), 121–138.
- Lehmann, M. & Kopiez, R. (2011). Der Musikgeschmack im Grundschulalter – Neue Daten zur Hypothese der Offenohrigkeit. In W. Auhagen, C. Bullerjahn & H. Höge (Hrsg.), *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie* (Bd. 21, S. 30–55). Göttingen: Hogrefe.
- Massey, F.J. (1952). Distribution table for the deviation between two sample cumulations. *Annals of Mathematical Statistics* 23, 435–441.
- Quinlan, R. (1993). *C4.5: Programs for Machine Learning*. San Mateo, CA.: Morgan Kaufmann.
- Rinne, H. (2003). *Taschenbuch der Statistik*. Frankfurt am Main: Harri Deutsch.
- Shannon, C.E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, Illinois.
- Schellberg, G. & Gembris, H. (2003). Was Grundschulkindern (nicht) hören wollen. Eine neue Studie über Musikpräferenzen von Kindern der 1. bis 4. Klasse. *Musik in der Grundschule*, 4, 48–52.
- Tränkle, U. (1991). *Mathematische und statistische Methoden*. Münster: Aschendorff.