

# Dimensionen der Groove-Affordanz (DGA) beim Hören populärer Musik: Ein neues Messinstrument und eine vergleichende Validierungsstudie

Dimensions of Groove Affordances (DGA) When Listening to Popular Music: A New Measurement Instrument and a Comparative Validation Study

Steffen Lepa<sup>1</sup> , Luzie Ahrens<sup>1</sup>, Martin Pfeleiderer<sup>2</sup> 

[1] Fachgebiet Audiokommunikation, Technische Universität Berlin, Berlin, Deutschland. [2] Institut für Musikwissenschaft, Hochschule für Musik Franz Liszt Weimar, Weimar, Deutschland.

---

Jahrbuch Musikpsychologie, 2025, Vol. 33, Article e217, <https://doi.org/10.5964/jbdgm.217>

**Received:** 2025-03-12 • **Accepted:** 2025-08-28 • **Published (VoR):** 2025-10-09

**Reviewed by:** Toni Bechtold; Philippe Labonde.

**Corresponding Author:** Steffen Lepa, Fachgebiet Audiokommunikation (Sekt. EN-8), Technische Universität Berlin, Einsteinufer 17c, 10587 Berlin, Deutschland. E-mail: [steffen.lepa@tu-berlin.de](mailto:steffen.lepa@tu-berlin.de)

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



## Zusammenfassung

Beim Musikhören auftretende Bewegungsempfindungen und Bewegungsimpulse können ganz unterschiedliche Qualitäten besitzen und im Alltag bei hinreichendem Gefallen an der Musik und situativer Passung zu körperlichen Handlungen, z. B. Tanzbewegungen, führen. Um der qualitativen Vielfalt von musikinduzierten Bewegungsempfindungen zu begegnen, stellen wir mit den Dimensionen der Groove Affordanz (DGA) ein neues, theoretisch begründetes Messinstrument vor, welches auf einer Faktorenanalyse bereits erhobener Daten basiert und in der Lage ist, zwischen drei grundlegenden Dimensionen wahrgenommener rhythmischer Angebotsstrukturen beim Musikhören zu differenzieren: *Roll*, *Drive* und *Pulse*. Unter Nutzung digitaler Methoden der Datenerhebung (Webscraping, APIs, digitales Audiosampling und Online-Befragung) wird anschließend in Kombination mit Faktorenanalysen und Linear Mixed Models im Abgleich mit dem deutschsprachigen Experience of Groove Questionnaire (EGQ) untersucht, ob sich das neue Messinstrument auch an einem großen Stimuluspool von  $k = 448$  Popmusikstücken hinsichtlich Modellpassung, Reliabilität und Diskriminanzvalidität bewährt. Dazu werden nach Berechnung von Faktorenwerten aus  $n = 1775$  gesammelten Ratings von 178 Hörer\*innen die nach Genres aufgetretenen Unterschiede bezüglich der Bewegungsempfindungen beim Hören von Popsong-Exzerpten ausgewertet. Die Analyseergebnisse zeigen, dass das DGA auch bei einer großen Anzahl



musikalisch diverser Stimuli eine gute Modellpassung und Reliabilität aufweist und somit seinen Anspruch einlösen kann, unterschiedliche phänomenologische Aspekte von Bewegungsimpulsen bei vergleichsweise geringer Beeinflussung durch persönliches Gefallen abzubilden.

## Schlüsselwörter

Rhythmus, Entrainment, Faktorenanalyse, Musikgenres, Affordanzen, Körperbewegung

## Abstract

Movement sensations and impulses elicited by listening to music can vary greatly in quality and, in everyday contexts, may lead to physical actions such as dancing—provided the music is sufficiently enjoyable and the situation appropriate. To address the qualitative diversity of induced movement sensations, we introduce and methodologically validate the Dimensions of Groove Affordances (DGA), a new measurement instrument. Based on factor analyses of an existing dataset, the DGA distinguishes between three fundamental dimensions of perceived rhythmic structure during music listening: *roll*, *drive*, and *pulse*. Using digital methods of data collection (e.g., web scraping, APIs, digital audio sampling, and online survey) in combination with factor analyses and linear mixed models, we examine whether this new instrument demonstrates appropriate model fit, reliability, and discriminant validity—particularly in comparison to the German Experience of Groove Questionnaire (EGQ). The evaluation is conducted on a large stimulus pool of  $k = 448$  popular music excerpts. Factor scores were calculated from  $n = 1,775$  ratings provided by 178 participants. We then assessed genre-related differences in the movement sensations evoked by the short excerpts. Results indicate that the DGA exhibits good model fit and reliability across a musically diverse set of stimuli. It appears well-suited to capturing distinct phenomenological aspects of movement impulses with minimal bias resulting from individual musical preferences.

## Keywords

rhythm, entrainment, factor analysis, music genres, affordance, body movement

Einer der zentralen Aspekte beim Genuss populärer Musik ist die Groove-Erfahrung, sie bildet seit über zehn Jahren ein beliebtes Forschungsthema der Musikpsychologie. Analog zu einigen frühen empirischen, aber auch mehreren neueren Untersuchungen möchten wir im vorliegenden Beitrag dafür plädieren, in künftigen empirischen Studien ihren *enaktiven Charakter* stärker in den Blick zu nehmen, eine Sichtweise welche zunehmend auch durch musikpsychologische und neurowissenschaftliche Befunde gestützt wird. Dies impliziert, zusätzlich die beim Hören von Musik ausgelösten, qualitativ unterschiedlichen Arten von Bewegungsempfindungen und Bewegungsimpulsen systematisch zu erfassen. Diese Sichtweise begründen wir nachfolgend zunächst theoretisch sowie im Abgleich mit darauf bezogenen Forschungsergebnissen und stellen anschließend dazu ein neues Messinstrument vor.

## Groove in der Produktion und Rezeption populärer Musik

Groove wird aus Sicht von Musiker\*innen vor allem als ein Aspekt der gezielten Rhythmusgestaltung verstanden. Die Gestaltung zyklischer Patternstrukturen oder Grooves gilt besonders in afroamerikanisch geprägten populären Musikgenres (Jazz, Latin, Funk, R'n'B, HipHop, Reggae, Dub, Elektronik) als elementares Stilmittel (Pfeleiderer, 2002). Die sinnliche Erfahrung dieser Groovestrukturen wird manchmal (verkürzend) mit einem Gefühl der Tanzbarkeit gleichgesetzt. In diesem Sinne rangiert Groove in journalistischen und Alltagsdiskursen über populäre Musik häufig auch als wichtiger Faktor für die Popularität und den Erfolg entsprechender Musikstücke. Ein ganze Reihe von musikalisch-akustischen Faktoren werden in der Wissenschaft als ausschlaggebend für Groove-Erfahrungen diskutiert (Etani et al., 2024), so die Synkopierung, die Deutlichkeit des Grundschlags oder Beats (beat salience), Tempo und Ereignisdichte, Mikrotiming und dessen Abhängigkeit von Attack-Hüllkurven der Onsets sowie der Anteil von Bassfrequenzen am Gesamtklang und die Einfachheit der harmonischen Gestaltung.

Verkompliziert wird die empirische Groove-Forschung nun dadurch, dass Groove ähnlich wie von Musik ausgedrückte Emotionen und Bedeutungen schon allein alltagsprachlich sowohl (1) als Eigenschaft der organisierten, klingenden bewegten Formen als auch (2) als ihre direkte Wirkung auf die Rezipient\*innen sowie (3) als eine Form verkörperten musikalischen Handelns aufgefasst werden kann. Vor allem die beiden letztgenannten Sichtweisen auf Groove und ihre Implikationen sind Thema empirischer musikpsychologischer Forschung, welche typischerweise Fragebögen verwendet, wenngleich es auch einige neuere Ansätze für nicht-reaktive Messungen gibt (Etani et al., 2024).

## Groove-Erfahrung als Effekt akustisch-musikalischer Faktoren

Eine Vielzahl von Arbeiten der empirischen Musikpsychologie untersuchen die Groove-Erfahrung experimentell als einen direkten Effekt der in den Stimuli enthaltenen, oben genannten musikalischen Merkmale. In den letzten Jahren wurde dazu vor allem der Experience of Groove Questionnaire (EGQ; Senn et al., 2020) eingesetzt. Dieser wurde erfolgreich auf Reliabilität geprüft und inzwischen auch ins Deutsche übersetzt (Düvel et al., 2021). Er läuft auf die Messung der beiden als unabhängig konzipierten Dimensionen *Urge to move* und *Pleasure* hinaus. Schon in den Entwicklungsstudien zeigte sich allerdings eine sehr hohe Korrelation zwischen diesen beiden Faktoren im Messmodell der geschätzten konfirmatorischen Faktorenanalyse (0.78 englische Version, 0.90 deutsche Version). Auch in neueren Arbeiten tritt wiederholt das Problem der hohen Korrelation beider Subdimensionen auf, die zu nahezu identischen Messergebnissen auf beiden Faktoren führt und inzwischen auch teilweise problematisierend diskutiert wird (Kowalewski et al., 2020; Matthews et al., 2023; Pando-Naude et al., 2024; Romkey et al., 2025; Seeberg et al., 2024a, 2024b). Dieses in fast allen Studien unter Verwendung des

EGQ beobachtete Phänomen ist jedoch nicht nur ein messtheoretisches Problem, sondern verweist auch auf eine mögliche inhaltlich-theoretische Problemlage: Wenn Befunde der Groove-Forschung etwa darauf hinweisen, dass ein mittlerer Grad von Synkopierung als optimal groovend empfunden wird (Witek et al., 2014), könnte dies schlichtweg auch so interpretiert werden, dass dieser mittlere Grad an Synkopierung einfach am ehesten von den Hörer\*innen gemocht wird. Dieses Problem kann als erster Beleg dafür interpretiert werden, dass eine Fragebogen-Operationalisierung von Groove-Erfahrungen, die lediglich auf die subjektive Zustimmung zur empfundenen angenehmen Tanzlust abhebt, möglicherweise zu kurz greift. Es kann jedoch auch als Argument für die theoretische Validität eines alternativen enaktiven Verständnisses von Groove (Matyja & Schiavio, 2013) betrachtet werden, welches im nachfolgenden Absatz entfaltet wird. Allerdings: Da EGQ-Messungen mehrheitlich bislang in Experimental-Studien vorgenommen werden, die meist nur wenige, in der Regel vorwiegend rhythmische musikalische Stimuli wie Drum-Patterns darbieten, ist auch nicht auszuschließen, dass es sich bei der oft beobachteten hohen Korrelation lediglich um ein methodisches Artefakt eines zu kleinen und zu sehr auf Rhythmusforschung ausgelegten Stimuluspools handelt. Diesem Einwand versuchen wir in der vorliegenden Arbeit durch Anwendung des EGQ an einem großen Pool von populären Musikstücken unterschiedlicher Genres und mit Hörer\*innen unterschiedlichen Musikgeschmacks und unterschiedlicher Expertise zu begegnen.

## Groove-Erfahrung als Folge verkörpernten enaktiven musikalischen Handelns

Der vorherrschenden musikpsychologischen Sichtweise auf die Groove-Erfahrung liegt eine vielzitierte Definition von Janata et al. (2012) zugrunde: “[...] groove is that aspect of the music that induces a pleasant sense of wanting to move along with the music” (S. 56). Dieses Groove-Verständnis wird dann auch in vielen experimentellen Arbeiten der letzten Jahre kontinuierlich fortgeschrieben. Zwei Aspekte dieser Definition erscheinen uns jedoch zumindest kritikwürdig: Einerseits werden nicht alle von Musiker\*innen durch Instrumentalspiel oder Musikproduktion hervorgebrachte rhythmische Bewegungsimpulse von allen Hörer\*innen als angenehm empfunden. Wer einmal unfreiwillig in eine lautstarke Musikaufführung eines wenig präferierten Genres in einer Diskothek geraten ist, kann dies vermutlich anekdotisch bestätigen. Andererseits wird bei dieser Definition der aktive Beitrag von Hörer\*innen zum Entstehen des „Urge to move“ im Rahmen der Musikrezeption wohlmöglich nicht genügend berücksichtigt – und dafür gibt es sowohl theoretische als auch empirische Argumente.

Wie inzwischen auch von einigen Autorinnen aus der Musikpsychologie postuliert wird (Levitin et al., 2018; Witek, 2017; Duman et al., 2024), sollte die Entstehung von Groove-Empfindungen als Folge einer aktiven und (als angenehm empfundenen) verkörpernten Auseinandersetzung Musikhörender mit der bei der Wahrnehmung von Synkopen, Microtiming, Dynamik und Bass hervorgerufenen rhythmischen Angebots-

struktur der Musik aufgefasst werden, die wir im Folgenden als durch Musik ausgelöste Bewegungsempfindungen (englisch: feelings of motion) bezeichnen. Dass solche Bewegungsempfindungen existieren und relativ konsistent von Rezipient\*innen nicht nur passiv empfunden, sondern in vielen Fällen gleichsam innerlich oder äußerlich motorisch mitvollzogen werden, wird nicht nur seit langem von Vertreter\*innen der ökologischen Musikwahrnehmungstheorie unterstellt (Windsor & de Bézenac, 2012, Burger et al. 2013). Vielmehr zeigen dies auch empirische Studien mit non-reaktiven Verfahren wie Pupillometrie (Bowling et al., 2019), EMG (Todd & Cody, 2000) oder bildgebenden Verfahren der Neurowissenschaft (Matthews et al., 2020).

Während also davon auszugehen ist, dass stark rhythmische Musik durchaus in der Lage ist, bei ihren Hörer\*innen im Sinne eines unwillkürlichen Effekts eine Synchronisation von Körperfunktionen und Körperbewegungen im Sinne eines automatischen Entrainments auszulösen (Gonzalez-Sanchez et al., 2018; Kern & Lepa, 2014; Juslin, 2019), muss davon ausgegangen werden, dass die meisten Hörer\*innen diesen Impulsen im Alltag bei normalen Abhörlautstärken nicht notwendigerweise folgen oder sie sogar bewusst willentlich unterdrücken. Umgekehrt wird das aktive, top-down gesteuerte Enaktieren empfundener Bewegungsimpulse, der innerliche oder sogar äußerliche, verkörperte Mitvollzug der rhythmischen Strukturen inzwischen als maßgeblich für die Entstehung von Groove-Empfinden als auch für die begleitenden angenehmen Gefühle verstanden, wie Ergebnisse neurowissenschaftlicher Studien belegen (Matthews et al., 2020).

Der viel beschworene Genuss von Groove wäre also aus einer enaktiven Perspektive kein direkter Effekt der dargebotenen musikalischen Stimuli, sondern ein Nebenprodukt eines willentlich gesteuerten resonierenden Einschwingvorgang des Geist-Körpersystems der Hörer\*innen auf eine klingende, subjektiv vor dem Hintergrund der individuellen Erfahrungsgeschichte wahrgenommenen Angebotsstruktur im Sinne rhythmisch-klanglicher *Affordanzen* (Lepa, 2012), ein aktiv handelndes Sich-Einlassen (englisch: to enact), welches umso so wahrscheinlicher wird, je mehr die Musik persönlich gefällt und aktiv-nachvollziehend rezipiert wird und je mehr die Hörer\*innen eine Form situativer Passung empfinden, indem sie sich dem Groove hingeben. Sich eingrooven (Dudenredaktion, n.d.) beschreibt entsprechend dieser Sichtweise auch in der Jugend- und Alltagssprache den Umstand, dass Subjekte sich willentlich und aktiv in Resonanz zur rhythmisch-klanglichen Angebotsstruktur oder Affordanz von Musik begeben. Groove-Erfahrungen entstehen nach einer solchen enaktiven Sichtweise der Musikrezeption als Handlung (Matyja & Schiavio, 2013) also nicht einfach, weil Subjekte bestimmten musikalischen Stimuli mit einer bestimmten rhythmischen Struktur ausgesetzt werden, sondern weil diese die angebotenen Bewegungsimpulse aktiv innerlich und/oder äußerlich, also körperlich mitvollziehen (Roholt, 2014). Am Beispiel von Synkopen erläutert: “[...] syncopation opens up gaps in the music’s rhythmic surface, which listeners and dancers are invited to actively move into by synchronizing their movements

to the underlying beat. In this way, moderate levels of syncopation require listeners to *enact* the beat itself [...]” (Schiavio et al. 2024, S. 6). Dies wird bei musikkaffinen Menschen vermutlich häufig der Fall sein, und vor allem dann, wenn sie vor dem Hintergrund ihrer persönlichen, genrebezogenen Expertise einen geeigneten Zugang zu der spezifischen Angebotsstruktur des dargebotenen Rhythmus finden (Danielsen et al., 2022).

Die Plausibilität dieser Argumentation erhöht sich durch einige empirische Befunde aus der Musikpsychologie, die sämtlich darauf hindeuten, dass individuelle Faktoren wie Musikgeschmack, Vertrautheit und Expertise das Ausmaß der Groove-Erfahrung überraschend stark prägen.

## Individuelle Unterschiede bei der Intensität von Groove-Empfindungen

Matthews et al. (2019) konnten etwa zeigen, dass Gefallen an der Musik zu großen Teilen als Mediator zwischen der Komplexität der dargebotenen Musik und dem empfundenen Urge to move fungiert – ohne Gefallen und Affinität zur dargebotenen Musik würde demnach kaum ein Tanzdrang entstehen. Ähnlich demonstrierten Kowalewski et al. (2020), dass persönliche Sympathie für die Musiker\*innen bei identischen Musikstimuli zu signifikant stärkeren Groove-Empfindungen führt. Befunde von O’Connell et al. (2022) konnten ferner belegen, dass sich musikalische Expertise, konkret Musical Training und Perceptual Abilities (Müllensiefen et al., 2014), sowie das Ausmaß persönlicher Tanzerfahrungen (Rose et al., 2022) signifikant auf die Sensibilität für die empfundene Intensität des Groove-Charakters der dargebotenen Musiktiteln auswirken. Duman et al. (2024) zeigten mit einer interpretativen Analyse von Freitextangaben zu Groove-Erfahrungen, die im Rahmen einer Online-Befragung gewonnen wurden, dass Informant\*innen vor allem dann Groove erleben, wenn sie eine Beziehung zu den Künstler\*innen haben, die gespielte Musik mögen oder deren wahrgenommener emotionaler Ausdruck ihrer aktuellen Stimmung entspricht. Senn et al.’s (2018) Studienergebnisse belegten schließlich, dass persönliches Gefallen am Musikgenre und Vertrautheit eines Stimulus gemeinsam bereits 15% der Varianz in empfundener Groove-Intensität erklären, während alle dort geprüften Modelle mit ausschließlich musikalisch-rhythmischen Faktoren maximal 2% aufklären konnten.

## Die empirische Erfassung wahrgenommener rhythmischer Affordanzen

Zusammengenommen weisen diese Befunde darauf hin, dass Groove-Empfindungen zu einem nicht zu vernachlässigenden Anteil auf persönlichem Gefallen, expertisebedingtem Zugang sowie der situativ empfundenen Passung der jeweils dargebotenen Musik beruhen. Hieraus ziehen wir zwei Schlussfolgerungen, die direkt zum empirischen Teil dieser Arbeit hinführen:

1. Wenn das Mögen (als zentrale Facette von *Pleasure*) der dargebotenen Musik gleichzeitig eine notwendige Bedingung für das Entstehen von *Groove*-Erfahrungen (*Urge to move*) wie ihre Folge ist, macht es wenig Sinn, es als eine unabhängige Dimension der *Groove*-Erfahrung zu begreifen und zu messen. Die in zahlreichen jüngeren Studien aufgetretenen extrem hohen Korrelationen zwischen *Pleasure* und *Urge to move* bestätigen diese Auffassung. Gleichzeitig scheint uns eine eindimensionale, vorrangig auf ‚Tanzlust‘ abzielende Erfassung von rhythmischen Erfahrungen mit Musik der Komplexität der Phänomene nicht gerecht zu werden. Dies wird in neueren Arbeiten der Musikpsychologie auch zunehmend reflektiert (z. B. [Senn, Bechtold, Hoese, et al., 2023](#)).
2. Es scheint uns von daher sinnvoll, im Sinne des auch in der Musikwissenschaft zunehmend verbreiteten *Affordanz*-Konzepts zusätzlich die individuell wahrgenommene musikalisch-rhythmische Angebotsstruktur von Musikbeispielen zu erheben. Dabei geht es uns nicht darum, deren klanglich-musikalischen Strukturen zu beschreiben, sondern vielmehr darum, die hieraus bei den jeweiligen Hörer\*innen entstehenden Bewegungsempfindungen zu erfassen, unabhängig davon, ob Hörer\*innen dabei ein angenehmes Bedürfnis zu Tanzen verspüren und/oder sich überhaupt darauf enaktiv einlassen können. Eine solche Messung käme dem analytischen Ansinnen entgegen zu verstehen, welche musikalisch-rhythmischen Hinweisreize potenziell, also z. B. bei hinreichender Lautstärke sowie situativer Passung zwischen Hörer\*innen und Kontexten, einen angenehmen Bewegungswunsch bzw. *Groove*-Erfahrungen auslösen können. Dies dürfte für zahlreiche angewandte Forschungsbereiche, etwa die Entwicklung von Verfahren des *Music Information Retrieval* (MIR), zur Vorhersage des Grades der Tanzbarkeit musikalischer Stimuli zielführender sein als ein offenbar viel zu stark von individuellem Geschmack und Hörexpertise abhängiges Maß.

Während in den frühen theoretischen Arbeiten von Charles [Keil \(1987\)](#) bereits der enaktive Aspekt des Prozessierens von rhythmischen Strukturen als Auslöser von *Groove*-Erfahrungen vorgedacht wurde, haben sich unseres Wissens bislang nur drei Arbeiten mit der empirischen Erfassung wahrgenommener rhythmischer Angebotsstrukturen via Fragebögen befasst: [Gabrielsson \(1973\)](#) untersuchte mit Adjektivlisten die Empfindung der Bewegungsimpulse dargebotener Rhythmen und fand dabei mehrere unterschiedliche Dimensionen. [Madison \(2006\)](#) führte eine ähnliche Untersuchung mit 64 achtsekündigen Musikausschnitten durch und ermittelte die vier Dimensionen *Regularity*, *Groove*, *Swing* und *Flow*; er verwendete dabei *Groove* zusätzlich als Oberbegriff für die untersuchten Phänomene. [Pfeleiderer \(2010\)](#) verwendete insgesamt 77 Adjektive zur Beurteilung mehrerer Musikbeispiele und identifizierte per Hauptkomponentenanalyse (PCA) zahlreiche Subdimensionen der Bewegungsempfindung. Diese Daten wurden in der vorliegenden Arbeit im Rahmen einer Sekundäranalyse unter Verwendung alternativer Analyseverfahren nun zu einem neuen Messinstrument weiterentwickelt, welches zentrale Dimensi-

onen von Bewegungsimpulsen erfasst, die kohärent von einer großen Anzahl von Hörer\*innen unterschiedlicher musikalischer Sozialisation wahrgenommen werden können: Die Dimensionen der Groove-Affordanz (DGA).

## Forschungsfragen der beiden empirischen Teilstudien

In der zuerst durchgeführten Entwicklungsstudie der DGA wurde mit Hilfe einer Sekundäranalyse der Originaldaten von [Pfleiderer \(2010\)](#) unter Verwendung verschiedener, dem Datensatz angemessenerer statistischer Verfahren (EFA statt PCA, oblique Rotation, robuster Schätzer, Paralleltest zur Faktorenzahlbestimmung, vgl. [Fabrigar et al., 1999](#)) ein deutschsprachiges standardisiertes Fragebogeninstrument zur Messung zentraler Dimensionen der wahrgenommenen Groove-Angebotsstruktur (Affordanz) konstruiert.

Daran anschließend wurden in einer ersten Validierungsstudie die beiden deutschsprachigen Messinstrumente EGQ und DGA an einem über eine Online-Befragung gewonnenen Datensatz und unter Verwendung eines großen Stimuluspools aktueller populärer Musik im Hinblick auf Messmodellpassung, Faktorreliabilität ([Cho, 2016](#)) und Diskriminanzvalidität ([Rönkkö & Cho, 2022](#)) der gemessenen Dimensionen überprüft und zusätzlich einer ersten Erprobung der Kriteriumsvalidität ([Raykov, 2012](#)) in Bezug auf Genreunterschiede unterzogen. Die Forschungsfragen lauteten dabei:

1. Lassen sich die Dimensionen der beiden deutschsprachigen Instrumente zur Messung der Groove-Erfahrung (EGQ, DGA) mit einem größeren Musikpool aktueller populärer Musik mit hinreichender Passung, Reliabilität und Diskriminanzvalidität reproduzieren?
2. Auf welchen Dimensionen der beiden Instrumente zeigen sich unter Berücksichtigung von Musikvorlieben und Musikexpertise Unterschiede zwischen Musik aus verschiedenen Genres aktueller populärer Musik?

Zum gewählten Vorgehen der initialen Validierungsstudie ist anzumerken, dass es in Bezug auf die inhaltliche Validierung von Messinstrumenten für die Musikpsychologie grundsätzlich immer einen Zielkonflikt zwischen Fragen interner und externer Validität gibt. Das gilt auch für die Messung von Groove-Empfindungen, wie sich an existierenden Arbeiten aufzeigen lässt: Kontrollierte Experimentalstudien mit eher wenigen, systematisch ausbalancierten und auf die strenge Prüfung rhythmuspsychologischer Fragestellungen ausgerichteten Stimuli-Sets können helfen zu testen, ob vermutete musikalische Einflussfaktoren tatsächlich verantwortlichen für gemessene Bewegungsempfindungen sind – dieser Überprüfung wird sich in Zukunft auch der DGA Fragebogen stellen müssen. Umgekehrt bedarf es aber unseres Erachtens auch ex-post-facto Studien mit großen, repräsentativen und damit typischerweise nicht ausbalancierten Korpora erfolgreicher populärer Musik, um zu demonstrieren, dass sich ein neues Messinstrument auch zur groben Unterscheidung rhythmischer Erfahrungen bei von vielen Menschen heute im Alltag gehörter Musik bewährt. [Janata et al. \(2012\)](#) verwendeten in dieser Weise etwa

148 Musikausschnitte populärer Musik, die größtenteils aus dem iTunes Music Store stammten, und konnten Unterschiede in der Groove-Wahrnehmung zwischen vier verschiedenen Genres ermitteln. Allerdings nahmen an dieser Studie nur 19 Studierende teil. [Senn et al. \(2021\)](#) werteten 208 Stimuli populärer Musik der vergangenen sechs Jahrzehnte bei 233 Teilnehmenden eines Online-Experiments unter Berücksichtigung der Stück-Vertrautheit und aus den Daten abgeleiteten Meta-Genre-Vorlieben aus und fanden signifikant höhere Groove-Burteilungen für Stimuli aus zu Funk und Pop zusammengefassten Meta-Genres gegenüber Songausschnitten, die eher dem Rock zuzurechnen waren. Jene Studie diente der vorliegenden als Vorbild, wobei wir speziell daran interessiert waren, herauszufinden, ob sich auch mit aktueller populärer Musik, einer deutlichen feinteileren, durch Dritte vorgegebenen Genreklassifikation und einer nicht-ausbalancierten Stimulusauswahl signifikante Unterschiede bei der Rhythmusempfindung aufzeigen lassen. Zur Sicherung der Validität der Ergebnisse bei diesen erschwerten Voraussetzungen verwendeten wir statistische Analysemethoden, die mit stark unterschiedlichen Teilstichproben gut umgehen können und ergänzten diese durch Simulationsstudien zu Teststärke und Schätzpräzision.

## Methoden – Entwicklungsstudie des DGA

### Stichprobe, Stimuli, Ablauf und Messungen

Für die Entwicklung des deutschsprachigen Dimensionen-der-Groove-Affordanz- Fragebogens (DGA) wurden die von [Pfeleiderer \(2010\)](#) erhobenen Rohdaten einer Sekundäranalyse unterzogen. Die enthaltenen  $n = 134$  Ratings stammen von 38 Studierenden der Musikwissenschaft in zwei Uni-Seminaren in Hamburg und Weimar, die insgesamt neun verschiedene Musiktitel hörten (vgl. [Tabelle 1](#) für eine Übersicht gültiger Bewertungen pro Titel) und nach jedem Stück den Charakter der bei ihnen während des Hörens ausgelösten Empfindungen anhand einer Item-Liste von 77 Adjektiven auf einer 6-stufigen Skala von *trifft überhaupt nicht zu* bis *trifft genau zu* per Fragebogen beurteilten (Items in [Tabelle 9](#) im Anhang). Es handelt sich um sechs sehr rhythmische Popmusiktitel und drei geloopte Breakbeatsamples, die aufgrund ihres groovigen Charakters sehr häufig in HipHop und elektronischer Tanzmusik eingesetzt werden ([Pfeleiderer, 2006](#), S. 325). Weitere Details der Datenerhebung und damaligen Datenanalyse lassen sich dem Open-Access Artikel von [Pfeleiderer \(2010\)](#) entnehmen.

**Tabelle 1***Datenbasis für die Entwicklung des DGA*

Stimulus	Gültige Ratings
King Tubby: „Natural Dub“ (Sanctuary Records, 1976)	25
Jalal: „Word of the Wise“ (On The One, 1997)	17
John Scofield: „Lazy“ (Blue Note, 2000)	13
Nightmares on Wax: „Dextrous“ (Warp Records, 1989)	15
Curtis Mayfield: „Trippin' Out“ (RSO, 1980)	15
King Floyd: „Groove Me“ (Atlantic, 1971)	20
„Apache“ Breakbeat	9
„Amen“ Breakbeat	10
„Funky Drummer“ Breakbeat	10

## Statistische Analyse

Zur Ermittlung der den Item-Ratings zugrundeliegenden Dimensionen wurde zunächst auf der Matrix aller Beurteilungen eine Explorative Faktorenanalyse (EFA) gerechnet. Dabei wurde der minimum-residuals-Ansatz mit obliquer Oblimin-Rotation gewählt, die Bestimmung der Faktorenzahl erfolgte anhand des Paralleltest-Kriteriums. Anschließend wurde mit Hilfe der  $> .6$  ladenden Items jedes identifizierten Faktors explorativ ein initiales Faktormessmodell mit Einfachstruktur konstruiert und mit Hilfe des R-Pakets *lavaan* (Rosseel, 2012) auf Fit geprüft. Dabei wurde der Robust Maximum Likelihood Schätzer (MLR) verwendet, aufgrund der Messwiederholungen cluster-robuste Standardfehler berechnet und zur Beurteilung des Modellfits auf die robusten Varianten der Fitindizes CFI (Comparative Fit Index), RMSEA (Root Mean Square Error of Approximation) und SRMR (Standardized Root Mean Squared Residual) zurückgegriffen (Savalei, 2018). Auf Basis von Modifikationsindizes wurde dieses Initialmodell durch Weglassen einzelner Items schrittweise optimiert, bis eine angemessene Passung zu den Daten und somit ein gutes CFA-Messmodell erreicht war. Um sicherzustellen, dass bei gegebener Beobachtungszahl trotz stark unterschiedlicher Häufigkeiten von Messwiederholungen bei den Befragten hinreichende Teststärke für unverzerrte Parameterschätzungen und Fit-Indizes des Messmodells vorlagen, wurde mit MPLUS eine Monte-Carlo-Simulation für konfirmatorische Faktorenanalysen (Muthén & Muthén, 2002) mit 1000 Wiederholungen unter Annahme von vier Items pro Faktor, Faktorladungen von  $.7$ , Faktorkorrelationen von  $.3$ , Residualstreuungen von  $.4$  und einer inner-Subjekt-SD der Items von  $.5$  gerechnet (entspricht hier einem mittleren ICC von  $.33$ ). Laut den Ergebnissen (vgl. digitaler Anhang) sind bei gegebener Stichprobenstruktur Coverage-Werte zwischen  $.89$  und  $.98$  sowie eine Teststärke zwischen  $.97$  und  $1$  zu erwarten. Die prognostizierten Schätzer für sämtliche beta-Koeffizienten und Standardfehler sind mit unter  $5\%$  höchstens minimal verzerrt und

die Streuung der prognostizierten Modellfitindizes überschritt lediglich für den SRMR mit  $M = .08$  und  $SD = .01$  leicht die üblichen Cut-Off-Werte (Hu & Bentler, 1999).

## Ergebnisse – Entwicklungsstudie des DGA

Die Stichprobe wies mit Blick auf einschlägige Simulationsstudien hinsichtlich ihrer Größe von  $n = 134$  die prinzipielle Eignung zur Entdeckung von mindestens 6 unterschiedlichen Faktoren auf (Fabrigar et al., 1999), keine der eingebrachten Variablen hatte zudem MSA-Werte kleiner als .50. Der Paralleltest wies auf eine optimale Repräsentation der in Beurteilungsmatrix enthaltenen Itemvarianzen durch vier oder weniger Faktoren hin. Die nach der oblimin-Rotation erhaltene Ladungsmatrix mit vier Faktoren, die zusammen ungefähr 50% der Itemvarianzen aufklären, findet sich als [Tabelle 9](#) im Anhang.

Der erste Faktor (23% der Itemvarianzen) wurde als *Drive* bezeichnet, weil die auf ihm hochladenden Adjektive („hämmernd“, „heftig“, „ruhelos“, „gehetzt“) das Phänomen eines Durch-den-Rhythmus-angetrieben-Werdens beschreiben, das auch von den Informant\*innen von Duman et al. (2024, S. 104ff.) betont wurde, außerdem lädt das Item „vortriebstreibend“ einzig auf diesem Faktor. Der zweite Faktor (9% der Itemvarianzen) wurde als *Roll* bezeichnet, weil die Adjektive „bewegungsanregend“, „schwungvoll“, „lebendig“ und „vital“ eine leichte, angenehm lebendige Bewegungsempfindung zu bezeichnen scheinen, wie sie zum Beispiel auch mit „flowing“, „swinging“ in der Studie von Madison (2006) auftrat. Der dritte Faktor (9% der Itemvarianzen) wurde *Pulse* genannt, weil die auf ihm hoch ladenden Adjektive („einheitlich“, „beständig“, „regelmäßig“, „geordnet“) eine eher monotone, sich wiederholende rhythmische Anregung des Körpers durch die Musik beschreiben, die auch bei Duman et al. (2024, S. 105) von einigen Befragten thematisiert wurde. Der vierte Faktor (8% der Itemvarianzen) wurde als *Play* bezeichnet, weil er mit den Adjektiven „freudig“, „verspielt“, „feierlich“, und „schwer“ (negativ ladend) die Dimension der Freude und Verspieltheit der erfahrenen Bewegungsempfindung zu erfassen scheint. Das aus dieser EFA-Lösung konstruierte initiale CFA-Messmodell wies mit  $n = 134$ ,  $X^2 = 1516.74$ ,  $df = 658$ ,  $p < .001$ ,  $CFI = .677$ ,  $SRMR = .117$ ,  $RMSEA = .099$  erwartungsgemäß zunächst einen kaum akzeptablen Fit zu den Daten auf. Daher wurden nachfolgend schrittweise Items eliminiert, die entweder zu gering oder in vergleichbarer Höhe auf mehreren Faktoren luden. Dieser Optimierung fiel schließlich der gesamte vierte Faktor zum Opfer, da seine Itemresidualvarianzen durchgehend zu stark mit den Itemresidualvarianzen des Faktors *Roll* korrelierten.

Das auf diesem Weg erreichte finale CFA-Messmodell mit den drei Dimensionen Drive, Roll und Pulse (vgl. [Tabelle 2](#)) weist trotz der vergleichsweise geringen Anzahl zugrundeliegender Beobachtungen einen guten Fit auf ( $n = 134$ ,  $X^2 = 65.24$ ,  $df = 51$ ,  $p = .087$ ,  $CFI = .974$ ,  $SRMR = .060$ ,  $RMSEA = .046$ , 95% CI [.001, .073]). Die modellimplizierten Faktorkorrelationen betragen  $\beta_{\text{RollDrive}} = .24$ ,  $\beta_{\text{DrivePulse}} = .03$ ,  $\beta_{\text{RollPulse}} = -.05$ .

**Tabelle 2***Finales CFA-Messmodell für Dimensionen der Groove-Angebotsstruktur*

Faktor	Item	<i>b</i>	<i>SE</i>	<i>Z</i>	<i>p</i>	stand. Ladung
Drive	gehetzt	1.03	0.10	10.72	< .001	0.83
	heftig	1.06	0.11	9.86	< .001	0.79
	ruhelos	0.93	0.11	8.54	< .001	0.70
	hämmernd	0.89	0.10	8.62	< .001	0.70
Roll	vital	1.01	0.09	10.71	< .001	0.83
	lebendig	0.84	0.09	9.65	< .001	0.77
	schwungvoll	0.93	0.10	9.17	< .001	0.74
	bewegungsanregend	0.81	0.10	7.95	< .001	0.66
Pulse	regelmäßig	0.69	0.08	8.63	< .001	0.75
	einheitlich	0.70	0.09	7.87	< .001	0.70
	beständig	0.58	0.08	7.36	< .001	0.69
	geordnet	0.59	0.090	6.58	< .001	0.60

## Methoden – Validierungsstudie (DGA und EGQ)

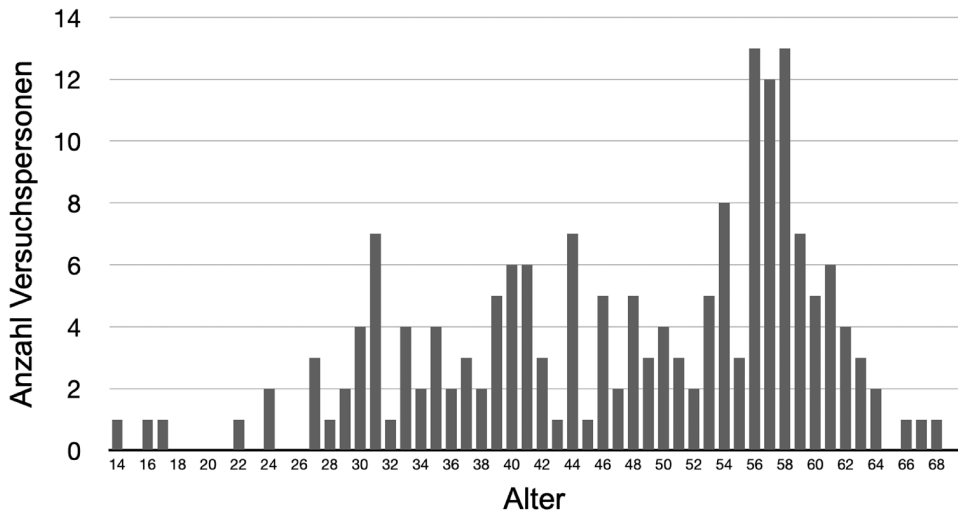
Zur ersten Validierung des neuen DGA-Messinstruments wurde im Frühjahr 2024 eine Online-Befragungsstudie durchgeführt, deren Ziel es war, das neue Instrument mit einer größeren Anzahl von Proband\*innen an einem möglichst großen und diversen Pool musikalischer Stimuli aus dem Bereich aktuell populärer Musik im Abgleich mit dem deutschen EGQ zu testen.

### Stichprobe

Die Stichprobe umfasste 178 Personen (113 weiblich, 60 männlich, 5 divers), mit einem biologischen Alter von  $M = 47.7$ ,  $SD = 11.7$ ,  $Min = 14$ ,  $Max = 68$ . [Abbildung 1](#) zeigt die Altersverteilung. Bei der Rekrutierung möglicher Teilnehmer\*innen wurden verschiedene Kanäle benutzt, und zwar E-Mail-Verteiler der Universität, persönliche Kontakte der Autor\*innen sowie verschiedene Social Network Communities.

**Abbildung 1**

Verteilung der Teilnehmer\*innen nach Alter

**Stimuli**

Die verwendete Musik stammte aus den *US-Billboard-Year-End-Hot-100-Single-Charts* von 2017 bis 2022. Dazu wurde eine Titelliste per Webscraping von <https://www.billboard.com/charts/year-end/> erstellt. Nach Entfernung von 57 Duplikaten verblieben 543 Songs, deren Spotify-URIs mit der Spotify-API ermittelt wurden, um von diesen mit Hilfe des Spotify-Web-Players und eines Python-Skriptes 30 Sekunden lange Exzerpte digital zu sampeln, die für die spätere Befragung ins MP3-Format enkodiert wurden. Anschließend wurde mittels der `discogs_client`-Python-Bibliothek (Joalla, 2024), die bei Discogs hinterlegten primären Genre-Informationen für alle Musiktitel ermittelt. Geling dies nicht, wurde der jeweilige Titel von der weiteren Analyse ausgeschlossen. Der finale Stimuluspool enthielt 450 Song-Exzerpte, die den Genres *Electronic* (66), *Folk, World & Country* (58), *Funk, Soul & R'n'B* (18), *HipHop* (183), *Latin* (6), *Pop* (93), *Reggae* (2) und *Rock* (24) durch die Nutzer\*innen von Discogs zugeordnet waren. Nach manueller Überprüfung der aufgelisteten Titel wurde nachfolgend *Country* für die Discogs-Kategorie „Folk, World & Country“ (Country-Songs dominierten eindeutig diese Kategorie) und *R'n'B* für die Kategorie „Funk, Soul & R'n'B“ (entsprechend der heute gängigen Genrebezeichnung) verwendet. Unter *Electronic* werden bei Discogs sowohl *Electronic Dance Music* als auch *Electronic Pop*, z. B. Songs von Billie Eilish, verstanden.

## Ablauf und Messungen

Die Teilnehmer\*innen gelangten über eine URL im Studienaufruf zu dem mit Limesurvey programmierten Online-Fragebogen. Nach explizitem Einverständnis mit der Datenschutzerklärung wurden sie um die Angabe von Geburtsjahr und Geschlecht und ihrer persönlichen Präferenz für jedes der elf im Stimuluspool vorkommenden Popmusik-Genres auf einer 7-stufigen Skala von *höre ich sehr ungerne* bis *höre ich sehr gerne* gebeten. Nachfolgend hatten sie auf einer 5-stufigen Ordinalskala anzugeben, wie häufig sie selbstgewählte Musik im Alltag hören und wie häufig sie sich körperlich im Alltag zu Musik bewegen (von *selten bis nie* bis *mehrmals täglich*).

Danach wurden die Items der Goldsmith-Musical-Sophistication-Index-Subskalen Musical Training und Perceptual Abilities (Schaal et al., 2014) mittels der vorgesehenen 7-stufigen Likert-Skalen vorgelegt. Anschließend erfolgte ein kurzer Test mit einem akustischen Beispielstimulus zur Sicherstellung einer hinreichenden Wiedergabelautstärke.

Im Hauptteil der Befragung wurden den Teilnehmer\*innen auf zehn Seiten jeweils ein Audioplayer zum Abspielen eines zufällig aus dem Stimuluspool ausgewählten Musikexzerpts dargeboten, sowie der Itemkatalog des EGQ mitsamt Instruktion („Bitte geben Sie Ihre Zustimmung zu den nachfolgenden Aussagen bezüglich des Musikbeispiels an.“) und der Itemkatalog des DGA mitsamt Instruktion („Beurteilen Sie, wie sehr die folgenden Adjektive zu den Bewegungs-Empfindungen passen, welche das Musikbeispiel bei Ihnen während des Hörens ausgelöst hat.“). Der Wortlaut der Items findet sich in [Abbildungen 2](#) und [3](#). Die Reihenfolge der Itemdarbietung war dabei für jede Versuchsperson unterschiedlich randomisiert. Abschließend wurden die Teilnehmenden dichotom gefragt, ob sie den jeweiligen Song bereits kannten.

Um die Datenqualität im Kontext einer Online-Befragung zu sichern, waren sämtliche Fragebogenangaben verpflichtend – die Befragung konnte somit nur abgeschlossen werden, wenn sämtliche Musikexzerpte abgespielt und sämtliche Fragen auch beantwortet worden waren. Ein im Browser lokal gespeicherter Cookie sorgte dafür, dass auch bei etwaigem Browserschließen oder Browserabsturz die Befragung später an der richtigen Stelle fortgesetzt werden konnte. Die Bearbeitung des gesamten Online-Fragebogens nahm im Median 19 Minuten in Anspruch.

## Statistische Analysen

Der Fragebogen stand vier Wochen lang online. Danach erfolgte eine deskriptive Analyse des Datensatzes. Songs, die zufallsbedingt weniger als drei Beurteilungen erhalten hatten, wurden von der weiteren Analyse ausgeschlossen, gleiches galt für Genres mit weniger als sechs beurteilten Titeln.

Zur Beantwortung der ersten Forschungsfrage wurden dann zwei konfirmatorische Faktorenanalysen (CFA) durchgeführt, eine für das EGQ und eine für die DGA. Hierfür

wurde das R-Paket *lavaan* (Rosseel, 2012) in Kombination mit *semTools* (Jorgensen et al., 2022) verwendet. Dabei wurde das Robust-Maximum-Likelihood-Schätzverfahren (MLR) eingesetzt, aufgrund der 10-fachen Messwiederholung cluster-robuste Standardfehler berechnet und zur Beurteilung des CFA-Modellfits auf die robusten Varianten der Fitindizes CFI, RMSEA und SRMR zurückgegriffen (Savalei, 2018). Um sicherzustellen, dass für beide Messmodelle bei gegebener Anzahl von Beobachtungen trotz der Messwiederholungen innerhalb der Befragten unverzerrte Parameterschätzungen und Fit-Indizes der Messmodelle geschätzt würden, wurde für beide Messmodelle jeweils eine Monte-Carlo-Simulation für CFA (Muthén & Muthén, 2002) mit 1000 Wiederholungen auf Basis der mutmaßlichsten Populationswerte für EGQ (Düvel et al., 2021) und DGA (Studie 1) gerechnet. Programmiercode und detaillierte Ergebnisse dieser Analyse sind im digitalen Anhang zu finden. Laut den Ergebnissen (vgl. digitaler Anhang) sind bei der gegebenen Stichprobenstruktur für alle Modellschätzer Coverage-Werte zwischen .93 und .96 sowie durchgehend eine Teststärke von 1 gegeben, die prognostizierten Schätzer für sämtliche beta-Koeffizienten und Standardfehler sind unter 2% verzerrt und die prognostizierte Streuung der zu erwartenden Modellfitindizes überschritt keinen der üblichen Cut-Off-Werte (Hu & Bentler, 1999).

Auf den somit gültigen CFA-Ergebnissen aufbauend wurden für beide Instrumente Faktorreliabilitäten (McDonald, 1999) berechnet und regressionsbasiert empirische Faktorwerte für alle Beurteilungen der Teilnehmer\*innen geschätzt. Außerdem wurden zu Vergleichszwecken auch einfache Mittelwertindizes für beide Konstrukte berechnet, da diese Operationalisierung in manchen Studien mit dem EGQ auch verwendet wird. Anschließend wurden bivariate Korrelationen der geschätzten Faktorwerte beider Instrumente sowie den Mittelwertindizes ermittelt. Ferner wurden auch für die Goldsmith-Musical-Sophistication-Index-Subskalen Perceptual Abilities und Musical Training die Mittelwertindizes berechnet. Zusätzlich wurde zu jeder Song-Beurteilung jeder Versuchsperson eine Kontrollvariable (*personal\_genre\_liking*) berechnet, welche das Präferenz-Rating der Versuchsperson für das jeweils in einem Trial dargebotene Genre enthält (ähnlich des *Style Bias* bei Senn et al. 2018, jedoch hier auf Basis der durch die discogs-Nutzer\*innen kollektiv vergebenen Genrekategorien).

Zur Beantwortung der zweiten Forschungsfrage wurden anschließend zur Aufklärung der Varianz in den fünf Faktorvariablen und zwei Mittelwertindizes nach Zentrierung aller metrischen Variablen jeweils ein Linear Mixed Model (LMM) mit Hilfe von *lmer* (Bates et al., 2015) berechnet. Als einziger auf die Forschungsfrage abzielender Prädiktor ging die Genrezugehörigkeit des dargebotenen Stimulus als Nominalvariable mit orthogonaler Effektkontrastcodierung in jedes der Modelle ein. Bei dieser Form von a-priori Kontrasten wird der Mittelwert jeder durch einen nominalen Faktor gebildeten Messwertgruppen im Vergleich zum Gesamtmittelwert aller Gruppen geprüft. Dabei bleibt eine einzige Kategorie des nominalen Faktors unberücksichtigt und geht nur in den Vergleichsmittelwert ein, hierfür wählten wir das Genre Pop aufgrund seines

tendenziell mehrdeutigen Charakters (Merlini, 2020). Ferner enthielt jedes Modell noch die sechs Kovariaten Alter, Perceptual Abilities, Musical Training, persönliche Hörintensität, persönliche Bewegungsintensität und personal genre liking. Als Clustervariable zur Berücksichtigung der Messwiederholungen in Form eines *Random Intercept* diente in den Modellen die Versuchspersonennummer. Aufgrund der unterschiedlichen Anzahl von Songs pro Genre wurden cluster-robuste Konfidenzintervalle berechnet und zusätzliche Simulationsstudien mit 1000 Wiederholungen für jedes der Modelle inklusive aller Kovariaten mit dem R-Paket *simr* (Green & MacLeod, 2016) durchgeführt, um zu überprüfen, ob die Teststärke für die Prüfung der Effektkontraste trotz der schwankenden Beobachtungszahlen in den verschiedenen Genres vergleichbar war – hierfür wurden kleine Effekte in Höhe von  $\omega^2 = .02$  unterstellt (Cohen, 1992). Die Simulation ergab für sämtliche zu prüfenden Mittelwertsunterschiede der genrebezogenen Ratings zum Gesamtmittelwert eine erwartete Teststärke von im Mittel 100% und der untere Rand des Konfidenzintervalls lag zwischen .994 und .996 (R-Code und Ergebnisdokumentation im digitalen Anhang). Darüber hinaus ist anzumerken, dass laut allgemeiner Simulationsstudien zu LMMs Koeffizientenschätzungen und Signifikanztests für feste Effekte ab  $k = 30$  Level-2-Einheiten nahezu unverzerrt sind, wobei die Anzahl der Messwiederholungen und der ICC demgegenüber kaum eine entscheidende Rolle spielen (Maas & Hox, 2005).

Nach erfolgter initialer Schätzung wurde für jedes Modell mit Hilfe eines LRT-Tests geprüft, ob der Random Intercept tatsächlich signifikant zur Modellverbesserung beitrug und ob die zusätzliche Modellierung von Random Slopes für die Genre-Effekte zu einer signifikanten Modellverbesserung führen würde. Für alle Modelle wurden anschließend ICC sowie marginales und konditionales  $R^2$  geschätzt und die Normalverteilung der Modellresiduen durch Inspektion von Quantil-Quantil-Plots sichergestellt. Schließlich wurden die Genre-Kontrastvariablen versuchsweise aus allen Modellen entfernt, um eine  $\Delta R^2_{\text{marg}}$  Schätzung für den Gesamtbeitrag der Genres zur Varianzaufklärung zu realisieren.

## Ergebnisse – Validierungsstudie (EGQ und DGA)

### Stichprobenstruktur

Bei der deskriptiven Analyse der Daten fiel auf, dass die zwei im Stimuluspool enthaltenen Reggae-Tracks jeweils weniger als drei Beurteilungen aufwiesen. Daher wurden diese aus der Stichprobe ausgeschlossen. Den resultierenden Datensatz bezeichnen wir als *Pop-Groove-Ground-Truth-Dataset* (PGGT). Er enthält  $n = 1775$  gültige Beobachtungen, die von  $m = 178$  Befragten stammen und ist online für Sekundäranalysen interessierter Kolleg\*innen verfügbar. Die Befragten beurteilten  $k = 448$  Popsong-Exzerpte, von denen ihnen 68% bekannt waren, auf den Erhebungsinstrumenten EGQ und DGA. Nach Genres aufgeschlüsselt liegen für Pop 376, für Country 215, für Electronic 275, für R'n'B 73, für

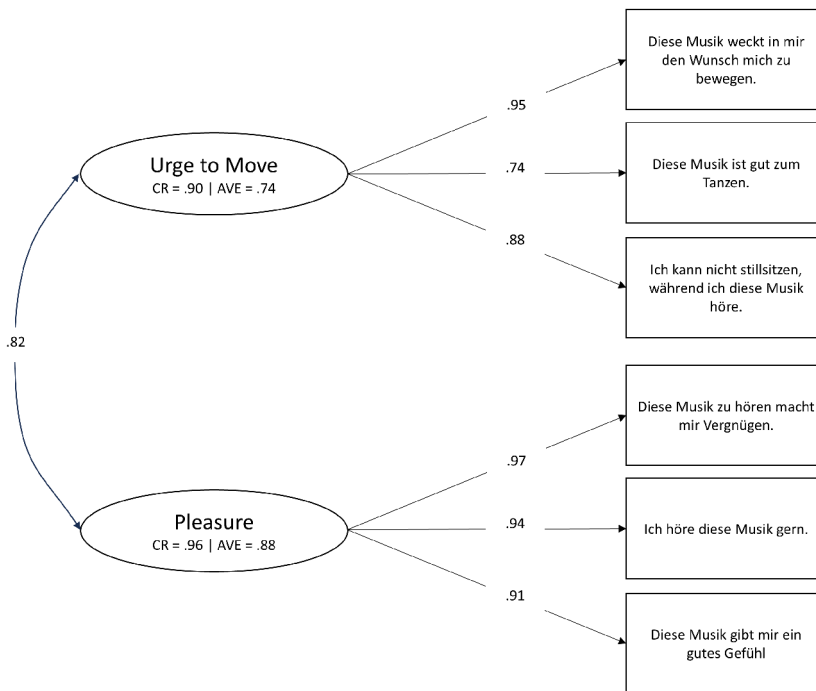
HipHop 721, für Latin 21, und für Rock 94 Beurteilungen vor. Die Ungleichverteilung ergibt sich einerseits aus den Fallzahlen der Genres im Stimuluspool, andererseits aus der zufälligen Selektion der Tracks aus dem Pool während der Online-Befragung.

## Güte der Messmodelle

Das geschätzte Messmodell des EGQ (Abbildung 2) erreicht nur einen befriedigenden Fit ( $n = 1775$ ,  $X^2 = 112.58$ ,  $df = 8$ ,  $p < .001$ ,  $CFI = .991$ ,  $SRMR = .019$ ,  $RMSEA = .082$ , 95%  $CI$  [.063, .103]) bei hervorragenden Faktorreliabilitäten ( $\omega_{Urge} = .90$ ,  $\omega_{Pleasure} = .96$ ), aber einer sehr hohen modellimplizierten Faktorinterkorrelation ( $\beta = .82$ ,  $CI_{Lower} = .79$ ,  $CI_{Upper} = .86$ ).

Abbildung 2

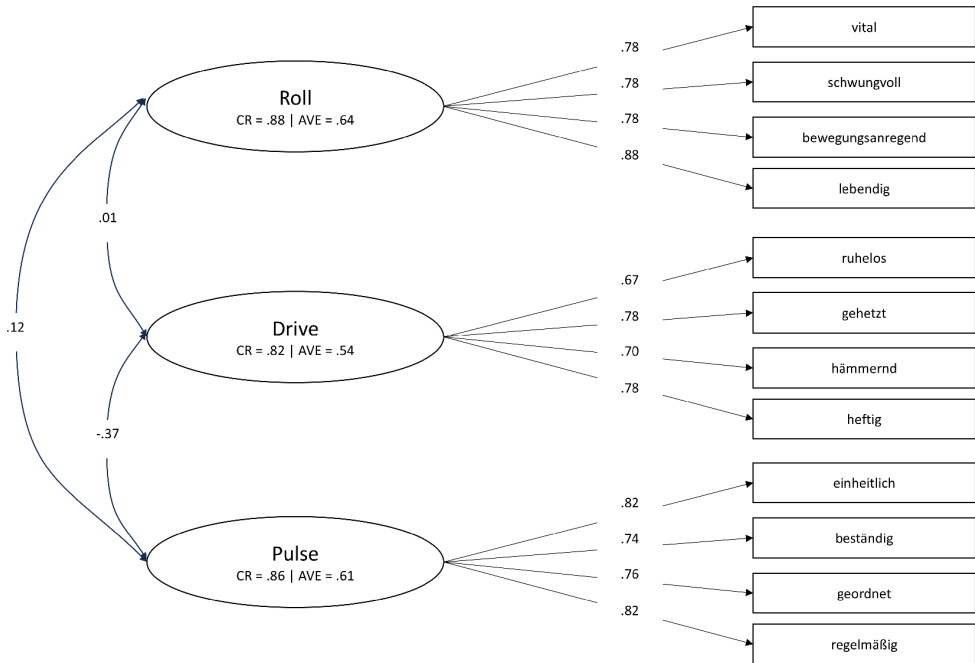
Pfaddiagramm der CFA zum Experience of Groove Questionnaire (EGQ) ( $n = 1775$ ,  $k = 178$ )



Das Messmodell des DGA (Abbildung 3) weist demgegenüber einen sehr guten Fit ( $X^2 = 394.86$ ,  $df = 51$ ,  $p < .001$ ,  $CFI = .971$ ,  $SRMR = .041$ ,  $RMSEA = .056$ , 95%  $CI$  [.047, .065]), sehr gute Reliabilitätswerte ( $\omega_{Roll} = .88$ ,  $\omega_{Drive} = .82$ ,  $\omega_{Pulse} = .86$ ) und nur geringe modellimplizierte Faktorinterkorrelationen ( $\beta_{RD} = 0$ ,  $\beta_{PR} = .12$ ,  $\beta_{PD} = -.37$ ) auf.

**Abbildung 3**

Pfaddiagramm der CFA zu Dimensionen der Groove Affordanz (DGA) ( $n = 1775$ ,  $k = 178$ )

**Bivariate Korrelation der Faktorwerte**

Analog zu den modellimplizierten Faktorkorrelationen fielen auch die empirischen Korrelationen der beiden regressionsbasiert geschätzten Faktorwerte des EGQ mit  $r = .86$  sehr hoch aus. Darüber hinaus zeigten sich ebenfalls hohe Korrelationen beider EGQ-Dimensionen mit dem Faktor Roll des DGA-Instruments. Die Mittelwertindizes des EGQ korrelierten mit 0.98 bzw. 1 mit den CFA-Konstruktvariablen (Tabelle 3).

**Tabelle 3**

*Bivariate Korrelationen der in der Validierungsstudie verglichenen Konstruktvariablen*

Variable	EGQ: CFA		EGQ: index		DGA: CFA	DGA: CFA	DGA: CFA
	Urge to move	EGQ: CFA Pleasure	Urge to move	EGQ: index Pleasure	Roll	Drive	Pulse
EGQ: CFA Urge to move	–						
EGQ: CFA Pleasure	.86	–					
EGQ: index Urge to move	.98	.80	–				
EGQ: index Pleasure	.84	1		–			
DGA: CFA Roll	.68	.61	.70	.60	–		
DGA: CFA Drive	-.11	-.21	-.08	-.23	-.00	–	
DGA: CFA Pulse	.14	.16	.13	.17	.14	-.43	–

## Unterschiede der Groove-Erfahrung bei Popmusik-Genres gemäß EGQ

Das LMM zum EGQ-Faktor Urge to move zeigt mit  $R^2_{\text{marg}} = .11$  und  $R^2_{\text{cond}} = .35$  eine gute Varianzaufklärung und einen  $ICC = .30$ , der auch zu signifikanter Verbesserung des Modells durch einen Random Intercept führte ( $LRT = 269.09$ ,  $df = 1$ ,  $p < .001$ ). Das Modell ließ sich durch Hinzunahme von Random Slopes nicht signifikant verbessern ( $LRT = 34.06$ ,  $df = 27$ ,  $p = .164$ ). Die geprüften Effektkontraste (Tabelle 4) und Effektplots (Abbildung 4) legen nahe, dass HipHop- und Country-Stücke gegenüber Titeln aus allen anderen Genres in signifikant unterdurchschnittlichem und Latin-Stücke in signifikant überdurchschnittlichem Maße Bewegungsdrang hervorriefen. Die Effektstärke der Genres im Modell beträgt ungefähr  $\Delta R^2_{\text{marg}} = .03$ . Zusätzliche negative signifikante Effekte auf den Urge to move gehen von der Häufigkeit des Musikhörens ( $\beta = -.09$ ,  $t(1760) = -2.07$ ,  $p = .038$ ) und positive von der Bewegungshäufigkeit zu Musik im Alltag aus ( $\beta = .15$ ,  $t(1760) = 3.25$ ,  $p = .001$ ). Außerdem gibt es einen hochsignifikanten Effekt der Vorliebe für das dargebotene Musik-Genre auf die Urge to move-Messung ( $\beta = .19$ ,  $t(1760) = 7.56$ ,  $p < .001$ ).

Tabelle 4

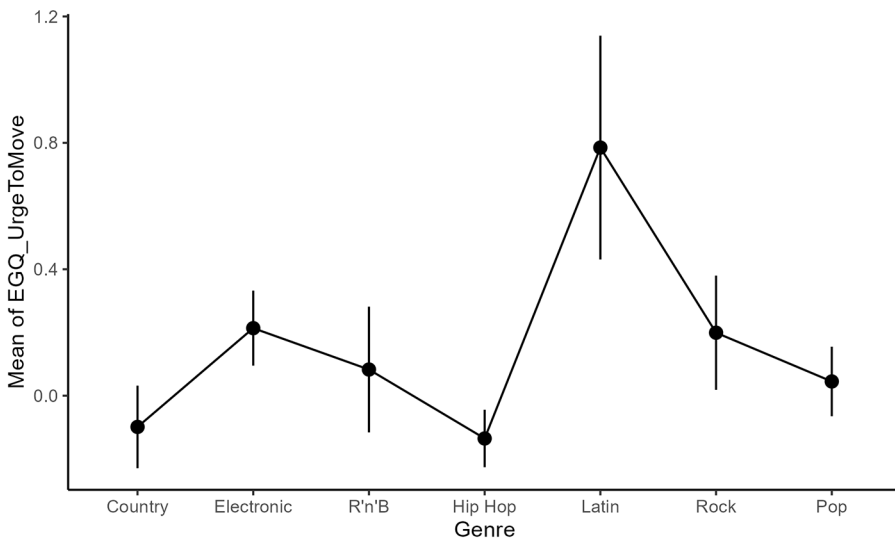
Geschätzte LMM-Effektkontraste für EGQ Urge to move: Genrebeurteilungen gegen Gesamtmittelwert

Songgenre	Vergleichswert	Differenz	SE	95% CI	$t$ ( $df = 1760$ )	$p$
Country	MW (alle Genres inkl. Pop)	-0.25	0.06	[-0.37, -0.14]	-4.35	< .001
Electronic	MW (alle Genres inkl. Pop)	0.06	0.05	[-0.05, 0.16]	1.08	.281
R'n'B	MW (alle Genres inkl. Pop)	-0.07	0.09	[-0.24, 0.10]	-0.84	.402
HipHop	MW (alle Genres inkl. Pop)	-0.29	0.04	[-0.37, -0.21]	-6.87	< .001
Latin	MW (alle Genres inkl. Pop)	0.63	0.15	[0.33, 0.93]	4.1	< .001
Rock	MW (alle Genres inkl. Pop)	0.04	0.08	[-0.11, 0.20]	0.54	.588

Anmerkung. Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten.

Abbildung 4

Geschätzte Genremittelwerte des LMM zu EGQ: Urge to move (Fehlerbalken repräsentieren 95% Konfidenzintervalle)



Für das LMM zum EGQ-Faktor Pleasure führt ein Random Intercept zur Verbesserung ( $LRT = 266.38$ ,  $df = 1$ ,  $p < .001$ ), es ließ sich durch Hinzunahme von Random Slopes zwar formal noch einmal signifikant verbessern ( $LRT = 61.32$ ,  $df = 27$ ,  $p < .001$ ), allerdings nur durch Inkaufnahme eines *singular fit* (instabile Schätzung aufgrund zu weniger Beobachtungen). Daher blieben wir bei der Random Intercept Modellierung. Diese zeigt im Ergebnis mit  $R^2_{\text{marg}} = .14$  und  $R^2_{\text{cond}} = .46$  eine gute Varianzaufklärung ( $ICC = .32$ ). Die

geprüften Effektkontraste (Tabelle 5) und der Mittelwertplot (Abbildung 5) zeigen, dass HipHop-Tracks gegenüber Titeln aus allen anderen Genres in signifikant unterdurchschnittlichem und Latin-Stücke in signifikant überdurchschnittlichem Grade gefielen. Die Effektstärke der Genres im Modell beträgt ungefähr  $\Delta R^2_{\text{marg}} = .03$ . Zusätzliche starke positive Effekte auf Pleasure gehen von der Kontrollvariable *personal\_genre\_liking* aus ( $\beta = .26$ ,  $t(1760) = 10.36$ ,  $p < .001$ ). Die Ergebnisse zu den über Mittelwertindizes geschätzten EQQ-Dimensionen fallen bei aufgrund der höheren Messfehler lediglich leicht reduzierten Effektstärken inhaltlich nahezu identisch aus. Sie sind im Anhang in Tabellen 10 und 11, sowie Abbildungen 9 und 10 zu finden.

**Tabelle 5**

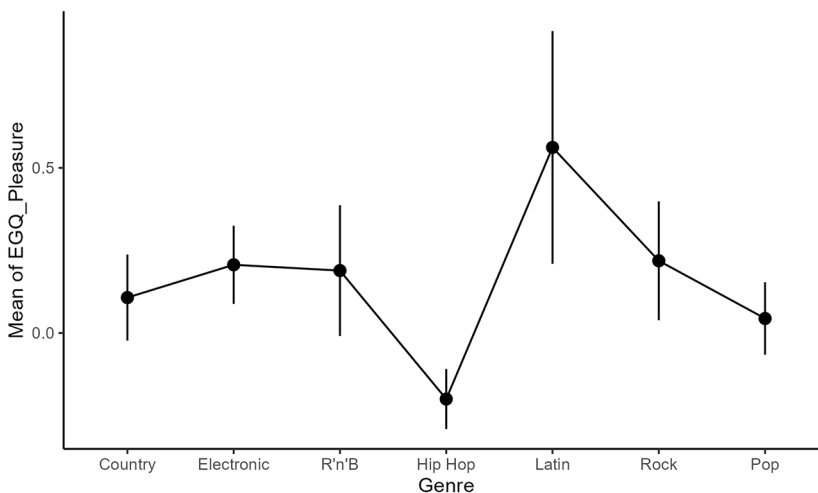
*Geschätzte LMM-Effektkontraste für EQQ Pleasure: Genrebeurteilungen gegen Gesamtmittelwert*

Songgenre	Vergleichswert	Differenz	SE	95% CI	$t(df = 1760)$	$p$
Country	MW (alle Genres inkl. Pop)	-0.05	0.06	[-0.17, 0.06]	-0.92	.356
Electronic	MW (alle Genres inkl. Pop)	0.05	0.05	[-0.06, 0.15]	0.85	.394
R'n'B	MW (alle Genres inkl. Pop)	0.03	0.09	[-0.14, 0.20]	0.32	.747
HipHop	MW (alle Genres inkl. Pop)	-0.36	0.04	[-0.44, -0.28]	-8.56	< .001
Latin	MW (alle Genres inkl. Pop)	0.4	0.15	[0.10, 0.70]	2.63	.009
Rock	MW (alle Genres inkl. Pop)	0.06	0.08	[-0.10, 0.21]	0.72	.469

*Anmerkung.* Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten.

**Abbildung 5**

*Geschätzte Genremittelwerte des LMM zu EQQ: Pleasure (Fehlerbalken repräsentieren 95% Konfidenzintervalle)*



## Unterschiede der Groove-Erfahrung bei Popmusik-Genres gemäß DGA

Das LMM zum DGA-Faktor Roll zeigt mit  $R^2_{\text{marg}} = .09$  und  $R^2_{\text{cond}} = .24$  eine gute Varianzaufklärung bei einem  $ICC = .17$  und signifikanter Modellverbesserung durch einen Random Intercept ( $LRT = 121.88$ ,  $df = 1$ ,  $p < .001$ ). Das Modell ließ sich durch Hinzunahme von Random Slopes nicht signifikant verbessern ( $LRT = 19.68$ ,  $df = 27$ ,  $p = .844$ ). Die geprüften Effektkontraste (Tabelle 6) und der Mittelwertplot (Abbildung 6) demonstrieren, dass HipHop-, R'n'B- und Country-Stücke gegenüber Songs aus allen anderen Genres signifikant unterdurchschnittliche Roll-Empfindungen hervorriefen, während Latin-, und Electronic-Stücke signifikant überdurchschnittliche Faktorwerte generierten. Die Effektstärke der Genres im Modell beträgt ungefähr  $\Delta R^2_{\text{marg}} = .05$ . Zusätzliche signifikante positive Effekte auf Roll gehen von den Kontrollvariablen Musical Training ( $\beta = -.09$ ,  $t(1760) = -2.10$ ,  $p = .036$ ) und Personal Genre Liking ( $\beta = .15$ ,  $t(1760) = 5.78$ ,  $p < .01$ ) aus.

**Tabelle 6**

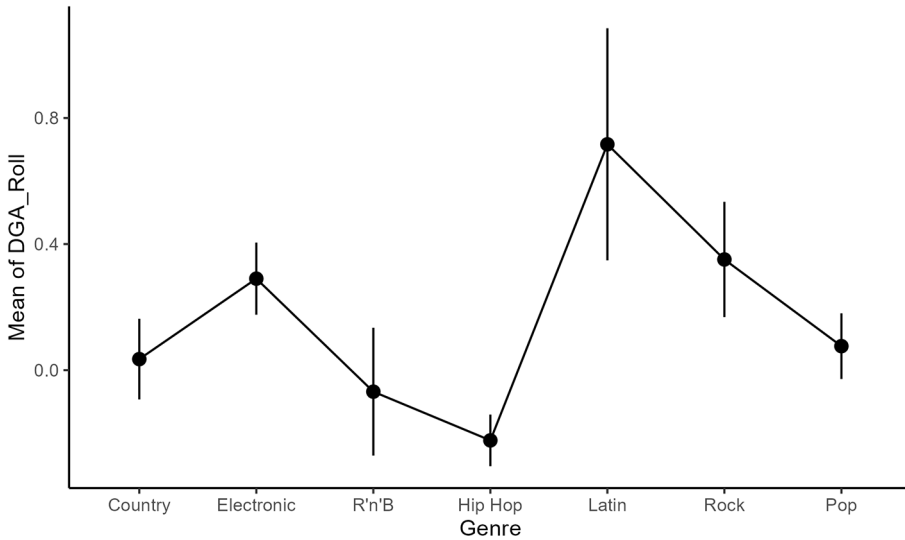
Geschätzte LMM-Effektkontraste für DGA Roll: Genrebeurteilungen gegen Gesamtmittelwert

Songgenre	Vergleichswert	Differenz	SE	95% CI	$t$ ( $df = 1760$ )	$p$
Country	MW (alle Genres inkl. Pop)	-0.13	0.06	[-0.25, -0.01]	-2.17	.030
Electronic	MW (alle Genres inkl. Pop)	0.12	0.06	[0.01, 0.23]	2.17	.030
R'n'B	MW (alle Genres inkl. Pop)	-0.24	0.09	[-0.42, -0.06]	-2.59	.010
HipHop	MW (alle Genres inkl. Pop)	-0.39	0.04	[-0.48, -0.30]	-8.8	< .001
Latin	MW (alle Genres inkl. Pop)	0.55	0.16	[0.23, 0.86]	3.4	< .001
Rock	MW (alle Genres inkl. Pop)	0.18	0.08	[0.02, 0.35]	2.18	.029

Anmerkung. Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten.

**Abbildung 6**

Geschätzte Genremittelwerte des LMM zu DGA: Roll (Fehlerbalken repräsentieren 95% Konfidenzintervalle)



Das LMM zum DGA-Faktor Drive zeigt mit  $R^2_{\text{marg}} = .10$  und  $R^2_{\text{cond}} = .43$  eine gute Varianzaufklärung bei einem  $ICC = .32$  und signifikantem Random Intercept ( $LRT = 453.45$ ,  $df = 1$ ,  $p < .001$ ). Das Modell ließ sich durch Hinzunahme von Random Slopes nicht signifikant verbessern ( $LRT = 39.51$ ,  $df = 27$ ,  $p = .057$ ). Die geprüften Effektkontraste (Tabelle 7) und der Mittelwertplot (Abbildung 7) zeigen, dass R'n'B- und Country-Stücke gegenüber Songs aus allen anderen Genres signifikant unterdurchschnittliche Drive-Empfindungen evozierten, während HipHop- und Electronic-Stücke signifikant überdurchschnittliche Faktorwerte auf dieser Dimension haben. Die Effektstärke der Genres im Modell beträgt ungefähr  $\Delta R^2_{\text{marg}} = .09$ . Keine der Kontrollvariablen übt einen direkten signifikanten Einfluss auf diese Faktordimension aus.

Tabelle 7

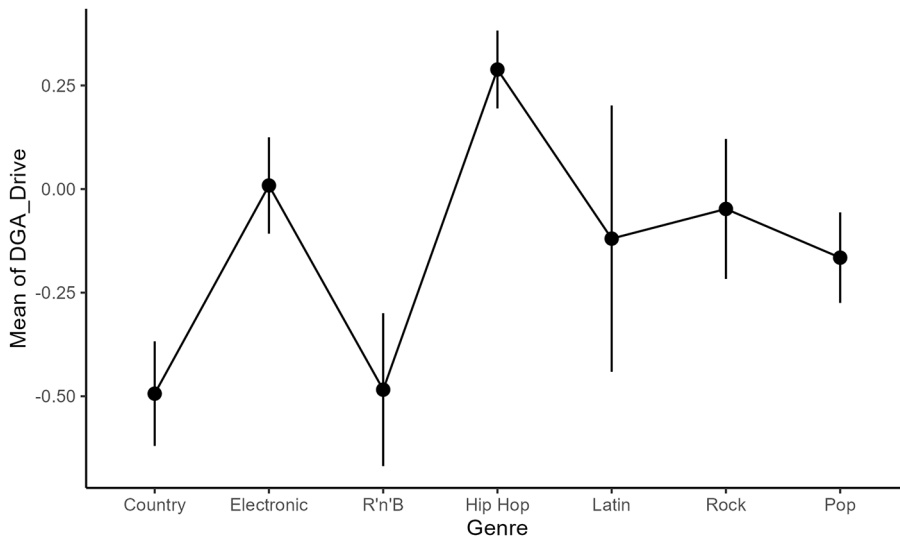
Geschätzte LMM-Effektkontraste für DGA Drive: Genrebeurteilungen gegen Gesamtmittelwert

Songgenre	Vergleichswert	Differenz	SE	95% CI	$t$ ( $df = 1760$ )	$p$
Country	MW (alle Genres inkl. Pop)	-0.35	0.05	[-0.45, -0.25]	-6.63	< .001
Electronic	MW (alle Genres inkl. Pop)	0.15	0.05	[0.06, 0.25]	3.19	.001
R'n'B	MW (alle Genres inkl. Pop)	-0.34	0.08	[-0.49, -0.19]	-4.33	< .001
HipHop	MW (alle Genres inkl. Pop)	0.43	0.04	[0.36, 0.51]	11.37	< .001
Latin	MW (alle Genres inkl. Pop)	0.03	0.14	[-0.25, 0.30]	0.18	.855
Rock	MW (alle Genres inkl. Pop)	0.1	0.07	[-0.04, 0.24]	1.35	.179

Anmerkung. Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten

Abbildung 7

Geschätzte Genremittelwerte des LMM zu DGA Drive (Fehlerbalken sind Fehlerbalken repräsentieren 95% Konfidenzintervalle)



Das LMM zum DGA-Faktor Pulse zeigt mit  $R^2_{\text{marg}} = .07$  und  $R^2_{\text{cond}} = .37$  eine gute Varianzaufklärung bei einem  $ICC = .34$  und signifikantem Random Intercept ( $LRT = 373.27$ ,  $df = 1$ ,  $p < .001$ ). Das Modell ließ sich durch Hinzunahme von Random Slopes nicht signifikant verbessern ( $LRT = 39.92$ ,  $df = 27$ ,  $p = .052$ ). Die geprüften Effektkontraste (Tabelle 8) und der Mittelwertplot (Abbildung 8) zeigen, dass HipHop-Tracks gegenüber Songs aus allen anderen Genres signifikant unterdurchschnittliche Pulse-Empfindungen evozierten, während Country-Stücke signifikant überdurchschnittliche Faktorwerte auf

dieser Dimension aufweisen. Die Effektstärke der Genres im Modell beträgt ungefähr  $\Delta R^2_{\text{marg}} = .03$ . Keine der Kontrollvariablen übt in diesem Fall einen direkten signifikanten Einfluss aus.

**Tabelle 8**

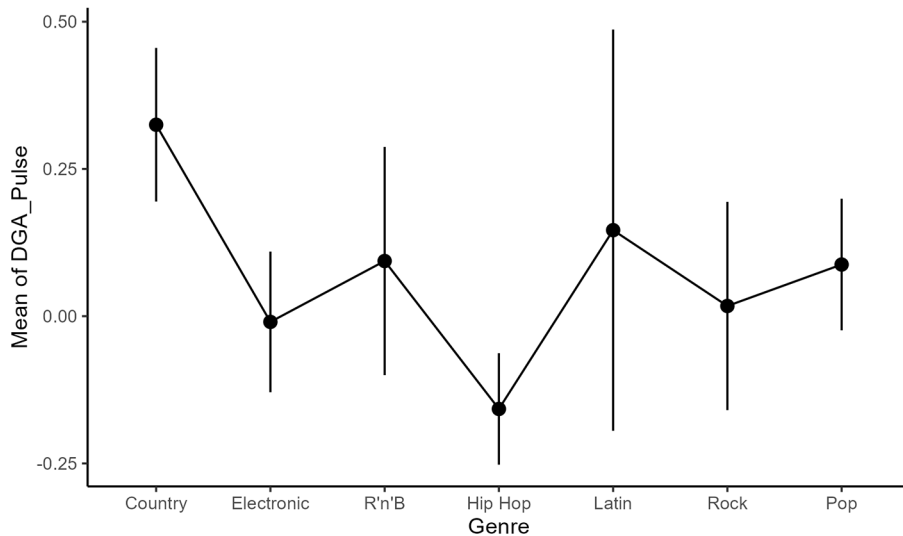
*Geschätzte LMM-Effektkontraste für DGA Pulse: Genrebeurteilungen gegen Gesamtmittelwert*

Songgenre	Vergleichswert	Differenz	SE	95% CI	$t$ ( $df = 1760$ )	$p$
Country	MW (alle Genres inkl. Pop)	0.25	0.06	[0.14, 0.36]	4.52	< .001
Electronic	MW (alle Genres inkl. Pop)	-0.08	0.05	[-0.18, 0.02]	-1.59	.112
R'n'B	MW (alle Genres inkl. Pop)	0.02	0.08	[-0.14, 0.19]	0.26	.793
HipHop	MW (alle Genres inkl. Pop)	-0.23	0.04	[-0.31, -0.15]	-5.65	< .001
Latin	MW (alle Genres inkl. Pop)	0.07	0.15	[-0.21, 0.36]	0.51	.613
Rock	MW (alle Genres inkl. Pop)	-0.05	0.08	[-0.20, 0.10]	-0.71	.477

*Anmerkung.* Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten.

**Abbildung 8**

*Geschätzte Genremittelwerte des LMM zu DGA: Pulse (Fehlerbalken sind Konfidenzintervalle)*



## Diskussion

### Ergebnisse der Entwicklungsstudie des DGA

Die in diesem Artikel dargestellte erste Teilstudie lieferte als Ergebnis eine schlüssige Faktorlösung zu den Dimensionen der Groove-Affordanz, die aufgrund der Verwendung strengerer Verfahren zur Faktorenselktion deutlich sparsamer ausfällt als die ursprüngliche Analyse derselben Daten von [Pfleiderer \(2010\)](#). Das resultierende Messinstrument in Form eines validierten deutschsprachigen Fragebogens aus zwölf Adjektiv-Items liefert eine einfache Methode, um mit Roll, Drive und Pulse drei grundlegend verschiedene Dimensionen rhythmischer Angebotsstrukturen im Sinne von Bewegungsempfindungen beim Hören von populärer Musik per Fragebogen zu messen. Semantisch weist das Ergebnis eine gewisse Überlappung mit den stärksten drei von [Madison \(2006\)](#) gewonnen Groove-Faktoren (Swing, Groove, Regularity) sowie mit den von [Senn, Bechtold, Hoesl, et al. \(2023\)](#) als Vorbedingung für Urge to move modellierten Konstrukten Temporal Regularity und Energetic Arousal auf (vgl. Diskussion weiter unten). Neben den guten Fit- und Reliabilitätswerten weist dies zusätzlich auf gute Inhaltsvalidität des DGA-Instruments in Bezug auf seine Zielsetzung hin, die wichtigsten Varianten beim Musikhören ausgelöster Bewegungsempfindungen besser differenzieren zu können. Einschränkend ist jedoch anzumerken, dass bei der Bildung dieses Datensatzes nur eine relativ kleine Musikauswahl und die Ratings weniger, eher musikkaffiner Hörer\*innen zurückgegriffen wurde. In Zukunft wäre es insofern sinnvoll und wichtig, das Vorgehen noch mit einer größeren Anzahl Hörer\*innen und mehr und diverseren Musikstücken zu wiederholen, und zugleich zu prüfen, ob Groove-Affordanzen gegebenenfalls sogar noch weitere Dimensionen aufweisen.

### Ergebnisse der Validierungsstudie (EGQ und DGA)

Die Ergebnisse der zweiten Teilstudie zeigen im Anschluss, dass sich beide untersuchten Fragebögen, der Experience of Groove Questionnaire in der deutschen Version und der neu entwickelte DGA-Fragebogen, prinzipiell erfolgreich auf eine große Auswahl aktueller Populärmusik und eine größere Schneeball-Stichprobe von Laien-Hörer\*innen anwenden lassen. Für den EGQ wurde trotz einer sehr großen Stichprobe nur ein befriedigender, für den DGA ein sehr guter Modellfit sowie in beiden Modellen sehr gute Faktorreliabilitäten erzielt. Allerdings verletzt die vom CFA-Modell implizierte Korrelation der beiden Faktoren des EGQ substanziell heute übliche Kriterien für Diskriminanzvalidität ([Rönkkö & Cho, 2022](#)). Weil beide Faktoren bereits im Messmodell zu stark miteinander korrelieren, können sie messtheoretisch kaum noch als distinkte Konstrukte betrachtet werden ([McDonald, 1985](#)), was dem selbst formulierten Anspruch des englischsprachigen Original-Messinstruments ([Senn et al., 2020](#)) zuwiderläuft, theoretisch und empirisch unabhängige Aspekte messen zu wollen.

In Bezug auf die Kriteriumsvalidität der beiden Instrumente ließen sich ferner die von Janata et al. (2012) und Senn et al. (2021) gefundenen Ergebnisse in Bezug auf Groove-Intensitäten unterschiedlicher Genres nicht wiederfinden: Weder erlebten die Hörer\*innen in der vorliegenden Studie bei dargebotenen HipHop- oder R'n'B-Stücken stärkere Urge to move-Erfahrungen als bei Titeln aus anderen Genres, noch generierten die gehörten Rock-Stücke geringere Empfindungen bezüglich dieses Konstrukts. Stattdessen wurde von den Angehörigen der vorliegenden, im Gegensatz zu den Vergleichsstudien etwa zehn Jahre älteren Stichprobe vor allem Titel aus dem Genre Latin als überdurchschnittlich groovig empfunden, während umgekehrt HipHop-Titel als vergleichsweise wenig bewegungsanregend beurteilt wurden. Sehr ähnliche Messergebnisse fanden sich, bedingt durch die hohe Korrelation der beiden Faktoren kaum verwunderlich, in Bezug auf den zweiten EGQ-Faktor Pleasure und werden darum hier nicht weiter diskutiert. Die ermittelten Genreunterschiede fielen beim EGQ insgesamt vergleichsweise klein aus (jeweils  $\Delta R^2_{\text{marg}} = .03$ ).

Vor dem Hintergrund der in der Einführung diskutierten möglichen Konfundierung des erfragten Bewegungsimpulses mit Gefallen, die sich in signifikanten Effekten von Genrevorlieben auf die Messung hier auch empirisch mittels der Kovariate Genre Liking belegen ließ, lässt sich vermuten, dass diese Divergenzen auf unterschiedliche musikalische Geschmäcker der hauptsächlich befragten Geburtskohorte zurückzuführen sein könnten: Wer sich von Musikstil und Künstler\*innen nicht angesprochen fühlt, wird auch deutlich weniger Urge to move empfinden.

Ein gegenüber den EGQ-Ergebnissen differenzierteres Bild ergaben die Messungen mit dem in dieser Arbeit entwickelten neuen DGA-Instrument: Die Messergebnisse zum DGA-Faktor Roll waren einerseits inhaltlich ziemlich ähnlich zu denen des EGQ-Faktors Urge to move (höhere Werte für Latin, niedrigere Werte für HipHop), wiesen dabei jedoch eine höhere Effektstärke auf ( $\Delta R^2_{\text{marg}} = .05$ ) und zeigten sich gleichzeitig empirisch deutlich weniger anfällig für eine Beeinflussung durch personenspezifische Variablen.

Ganz neue und mit bisherigen Studien kaum vergleichbare Messergebnisse zeigten sich beim DGA-Faktor Drive: Diese Facette von musikbedingten Bewegungsimpulsen erzeugte in unserer Studie die in Relation stärksten Unterschiede zwischen den Genres ( $\Delta R^2_{\text{marg}} = .10$ ) und wurde von den Befragten vor allem den Stilen HipHop und Electronic überdurchschnittlich stark zugerechnet, während sie bei R'n'B- und Country-Stücken als unterdurchschnittlich empfunden wurde. Dies entspricht zumindest auch anekdotisch den Empfindungen, die oftmals von Hörer\*innen jener Genres berichtet werden. Am schwächsten fielen die empfundenen Genre-Unterschiede in Bezug auf den DGA-Faktor Pulse aus ( $\Delta R^2_{\text{marg}} = .03$ ). Dennoch ließen sich auch diese statistisch belegen: So wurde eine pulsierende Bewegung überdurchschnittlich oft beim Hören von Country-Songs empfunden und unterdurchschnittlich oft bei HipHop-Tracks. Die Interpretation ist hier etwas schwieriger – allerdings scheint es uns durchaus plausibel, das Country-Songs aufgrund ihrer eingängigen rhythmischen Struktur, die typischerweise nicht auf Synkopen,

sondern auf stetig pulsierende Wiederholungen setzt, eine solche Bewegungsempfindung hervorbringen können.

## Allgemeine Diskussion

Mit dem DGA ist es uns gelungen, ein zum EGQ inhaltlich komplementäres Messinstrument zu entwickeln, welches jedoch zugleich konzeptionell eine grundlegend andere Zielsetzung verfolgt. Zwar kann die mit dem Begriff Groove assoziierte inhaltliche Facette des Bewegungsimpulses prinzipiell mit dem EGQ erfolgreich gemessen werden. Allerdings lässt dessen auch in unserer Erprobung aufscheinende inhaltliche Konfundierung mit Gefallen aus unserer Sicht die Erhebung des zweiten Faktors überflüssig erscheinen und stellt die konkrete Operationalisierung insgesamt in Frage. Dies wird in jüngeren Arbeiten zunehmend auch reflektiert. [Senn et al. \(2024\)](#) etwa scheinen Abstrahleffekte als ursächlich zu vermuten und ließen darum bei der Beurteilung von Drum-Stimuli die beiden Itemblöcke nur noch mit zeitlichem Abstand ausfüllen – ob dies das Problem tatsächlich auch bei der Anwendung auf größere Korpora populärer Musik lösen wird, bleibt jedoch abzuwarten. Wir raten daher Forscher\*innen, welche Fragebögen in der Groove-Forschung einsetzen, bis dahin mindestens zur zusätzlichen Verwendung des DGA, der – wie hier demonstriert – nicht so anfällig für den Einfluss von Gefallen bei der Messung ist. Außerdem liefert er zusätzliche unabhängige Messwerte für musikalisch induzierte Bewegungsempfindungen, die vom EGQ gar nicht erfasst werden, aber in Teilen der maßgeblichen Literatur, z. B. gerade auch bei qualitativ arbeitenden Befragungsstudien ([Hosken 2020](#), [Bechtold et al. 2023](#)) oder in der phänomenologisch ausgerichteten Monographie von [Roholt \(2014\)](#), thematisiert werden.

In unserem Theoriemodell werden durch Musik ausgelöste Bewegungsempfindungen als Affordanzen, also subjektiv wahrgenommene Handlungsoptionen konzipiert, die erst bei hinreichender Expertise, situativer Passung und Sympathie gegenüber Musik und Künstler\*innen zu einer Enaktierung von Groove und damit zur Empfindung eines körperlichen Bewegungsimpulses (Urge to move) führen. In ähnlicher Weise wird dies auch in zwei neueren Arbeiten von [Senn, Bechtold, Hoesl et al. \(2023\)](#) und [Senn, Bechtold, Jerjen et al. \(2023\)](#) postuliert und umgesetzt, welche die Wahrnehmung von Temporal Regularity und Energetic Arousal erfolgreich empirisch als Vorfaktoren für den empfundenen Urge to move belegten, den wiederum die Autoren daher inzwischen als „late response in a sequence of cognitive processes leading to body movement“ ([Senn, Bechtold, Jerjen, et al. 2023](#), S. 3) verstehen. Diese beiden Konstrukte weisen semantisch eine gewisse Verwandtschaft zu den von uns operationalisierten Faktoren Pulse, der Empfindung einer monoton pulsierenden rhythmischen Bewegung, und Drive, der Empfindung, durch den Rhythmus angetrieben zu werden, auf. Zusätzlich können wir aber mit der vorliegenden Arbeit zeigen, dass insbesondere Roll, die Empfindung rollender, schwungvoller Bewegungen, welche typisch für Latin Music zu sein scheint,

einen weiteren wichtigen Typ von Bewegungsimpulsen darstellen dürfte. Insofern sehen wir unsere vorliegenden Ergebnisse keineswegs im Widerspruch zum Konstrukt Urge to move und damit verbundener Forschung mit dem EQ, sondern eher als produktive Ergänzung. Ob man Aspekte der Bewegungsinduktion beim Hören, die nicht unbedingt als angenehm empfunden werden und nicht unbedingt zu einem Urge to move führen, konzeptionell als Teil der Groove-Erfahrung verstehen und analysieren möchte, wird von Forscher\*innen künftig auf Basis der zu untersuchenden Forschungsfrage, des untersuchten Korpus und des theoretischen Zugangs zum Thema Groove zu entscheiden sein. Wir geben noch einmal zu bedenken, dass auch alltagsanekdotisch das körperliche Erleben von Rhythmus und Entrainment-Phänomenen weder immer mit Bewegungslust noch mit Gefallen einhergehen muss, sondern durchaus auch negative, verstörende Gefühle umfassen kann, die jedoch wie alle vermeintlich ‚negativen‘ ästhetischen Erfahrungen ebenso als wertvoll und sinnstiftend empfunden werden können (Eden, 2020).

## Limitationen

Eine bereits erwähnte Limitation der dargestellten Entwicklungsstudie liegt in der sehr selektiven Auswahl von musikalischen Stimuli und Befragten, wodurch keineswegs auszuschließen ist, dass es womöglich noch mehr und andere Dimensionen der Groove-Affordanzen zu entdecken gäbe, als die drei von uns identifizierten.

Als Limitation der Validierungsstudie ist zu bedenken, dass die Genre-Zuordnungen von Songs bei Discogs wohlmöglich nicht hinreichend empirisch valide und reliabel sind, da sie eine *Folksonomy* bilden, also ähnlich wie Wikipedia-Einträge unsystematisch von verschiedenen Usern nach dem Mehrheitsprinzip vorgenommen werden. Eine stichprobenartige Überprüfung durch die Autoren ließ die vorgenommenen Kategorisierungen jedoch im Kern plausibel erscheinen. Grundsätzlich liegt ein Problem von poplärmusikalischen Genrebezeichnungen darin, dass sie historisch veränderliche soziale Konstrukte sind, die nicht nur auf musikalische Eigenschaften, sondern auch auf Hervorbringungs- und Rezeptionsmilieus abheben (Merlini, 2020). Darunter leiden allerdings auch die vergleichbaren Studien von Janata et al. (2012) und Senn et al. (2021). Ferner beschränkten wir uns bei der Stimulusauswahl in der Validierungsstudie vorrangig auf aktuelle und vor allem in den USA kommerziell erfolgreiche Popmusik der letzten Dekade, was letztlich auch die Vergleichbarkeit mit den beiden eben genannten Studien einschränkt – dabei mischen sich ganz unterschiedliche rhythmische Ausdrucksmittel und deren Intensitätsstärken innerhalb der dargebotenen Musikgenres. Wir glauben zwar nicht, dass die hier gewonnenen Ergebnisse zu Genreunterschieden den Anspruch erheben können, in irgendeiner Weise allgemeingültig zu sein. Wir verstehen sie jedoch, wie in der Einleitung dargelegt, als adäquates Mittel einer vergleichenden Prüfung der Diskriminationsfähigkeit und Kriteriumsvalidität beider Instrumente. Diese Einschränkung gilt allein schon deshalb, weil für einige Genres wie Latin einfach zu wenig Beobachtungen im Datensatz vorliegen. Die Ergebnisse zu den ebenfalls vergleichend untersuchten

psychometrischen Fragen (Modellpassung, Reliabilität, Diskriminanzvalidität) werden unseres Erachtens allerdings von diesem Problem nicht berührt, da die Instrumente sich invariant gegenüber der konkreten Stimulusauswahl verhalten sollten, wenn sie den Anspruch erheben wollen, prinzipiell zur differentiellen Messung der Wahrnehmung von Rhythmusphänomenen in aktueller populärer Musik geeignet zu sein.

## Ausblick

Insgesamt steht ungeachtet der dargestellten Ergebnisse eine überzeugende vergleichende Überprüfung der Kriteriumsvalidität beider in der vorliegenden Arbeit vorgestellten und analysierten deutschsprachigen Instrumente noch aus. Dazu sollten in Zukunft stärker kontrollierte Stimulussets, etwa aus Drum-Patterns, in eher experimentellen Studiendesigns eingesetzt werden, um zu untersuchen, welche spezifischen rhythmischen Gestaltungsmittel die unterschiedlichen durch das DGA gemessenen Bewegungsempfindungen jeweils hervorrufen. Ungeachtet dessen ließe sich ferner mit Recht kritisieren, dass die gewählte Methode der standardisierten Befragung gerade in Bezug auf den von uns theoretisch unterstellten enaktiven Charakter des Groove-Phänomens nicht optimal ist. Vordergründig günstiger wären daher unter Umständen Ansätze, musikalische Bewegungsinduktion über die Aufzeichnung von Körperbewegungen (Dotov et al., 2021) oder Psychophysiologie (Bowling et al., 2019) zu erheben. Umgekehrt können aber diese Verfahren ihre Validität nicht belegen, wenn nicht parallel auch die begleitenden subjektiven Empfindungen per Fragebogen erhoben werden. Ein vergleichbares Dilemma kennen wir bereits aus der Forschung zu Musik und Emotionen (Juslin & Sloboda, 2011). Ferner erscheint uns die Fragebogenmethode auch als alternativlos, wenn die gewonnenen Ergebnisse etwa mit Eigenschaften der wiedergegebenen Musikstücke mittels MIR-Verfahren in Verbindung gebracht werden sollen, wie etwa bei Düvel et al. (2022). Abschließend möchten wir noch einmal unser Argument bekräftigen, dass wegen der starken Erfahrungsgebundenheit und darum anzunehmenden sozialen Differenzierung von Rhythmuswahrnehmung neben Experimenten auch weiterhin Studien mit großen Korpora populärer Musik und entsprechenden Befragtenstichproben benötigt werden, die nur mit einem einfach zu handhabenden Fragebogen zu bewältigen sind. Nicht zuletzt auch aus diesem Grund möchten wir alle an Groove-Forschung Interessierten einladen, dabei mitzuhelfen, die gewonnene Datenbasis des Pop-Groove-Ground-Truth-Dataset in kollektiver Arbeit zu erweitern, und stellen dafür unseren bestehenden Datensatz und die im Code befindlichen Methoden zum Umgang mit komplexen nicht-ausbalancierten Stichproben interessierten Kolleg\*innen zur Weiterverwendung online zur Verfügung.

---

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

---

**Danksagung:** Wir bedanken uns herzlich bei Prof. Reinhard Kopiez (Hochschule für Musik, Theater und Medien Hannover) für seine Betreuungs- und Ko-Gutachtertätigkeit bei der dargestellten Validierungsstudie, die in einer früheren Fassung als Abschlussarbeit der Zweitautorin im Studiengang „Audiokommunikation und -technologie M.Sc.“ an der Technischen Universität Berlin eingereicht wurde.

---

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

---

**Beiträge der Autor\*innen:** Steffen Lepa – Konzeption | Methodik | Coding | Formale Analyse | Visualisierung | Supervision | Verfassen – Originalentwurf Entwicklungsstudie | Schreiben – Überarbeitung & Redaktion | Projektadministration. Luzie Ahrens – Konzeption | Methodik | Untersuchung | Coding | Datenkuratierung | Formale Analyse | Visualisierung | Verfassen – Originalentwurf Validierungsstudie | Visualisierung. Martin Pfeleiderer – Konzeption | Untersuchung | Datenkuratierung | Schreiben – Überarbeitung & Redaktion.

---

**Ethikerklärung:** Die dargestellte Entwicklungsstudie ist eine Sekundäranalyse bereits publizierter Daten (Pfeleiderer, 2010). Der empirische Teil der dargestellte Validierungsstudie wurde im Rahmen einer Masterarbeit an der Fakultät I für Geistes- und Bildungswissenschaften der Technischen Universität Berlin in Übereinstimmung mit den einschlägigen institutionellen und nationalen Richtlinien und Vorschriften (Föderation Deutscher Psychologenvereinigungen, 2022) sowie im Einklang mit den Prinzipien der Deklaration von Helsinki durchgeführt. Eine formale Genehmigung der Studie durch die Ethikkommission der Fakultät war nicht erforderlich, da alle geltenden Vorschriften für Befragungen im Rahmen von Masterarbeiten an der Fakultät eingehalten wurden. Alle Teilnehmenden gaben ihre schriftliche Einwilligung und wurden darüber informiert, dass sie das Recht haben, ihre Teilnahme an der Studie jederzeit ohne negative Konsequenzen abzubrechen.

---

**Datenverfügbarkeit:** Für diesen Artikel werden Datensätze und Skripte zur Verfügung gestellt (siehe Pfeleiderer & Lepa, 2025; Ahrens & Lepa, 2025; Lepa, 2025).

---

## Ergänzende Materialien

Für diesen Artikel sind folgende ergänzende Materialien verfügbar:

- Der Datensatz der Entwicklungsstudie (Daten aus Pfeleiderer, 2010) (siehe Pfeleiderer & Lepa, 2025)
- Der Datensatz der Validierungsstudie (PGGT - Pop-Groove-Ground-Truth-Dataset) (siehe Ahrens & Lepa, 2025)
- Die Skripte und Ergebnisse der zu den Faktorenanalysen und Linear Mixed Models dieser Arbeit durchgeführten Power-Simulationsstudien (siehe Lepa, 2025)

### Quellenverzeichnis der ergänzenden Materialien

Pfeleiderer, M. & Lepa, S. (2025). *DGE Dataset: Dimensionen der Groove-Erfahrung* [Datensatz]. Technische Universität Berlin. <https://doi.org/10.14279/depositonce-23174>

- Ahrens, L. & Lepa, S. (2025). *Pop Groove Ground-Truth Dataset (PGGT): Entwicklung einer Ground-Truth Datenbasis für die automatische Groove-Klassifikation von Popmusiktiteln* [Datensatz]. Technische Universität Berlin. <https://doi.org/10.14279/depositonce-23940>
- Lepa, S. (2025). *Power simulation study scripts and results for Lepa, Ahrens & Pfliederer (2025)* [Skripte, Ergebnisse]. Technische Universität Berlin. <https://doi.org/10.14279/depositonce-24335>

## Literatur

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bechtold, T. A., Kilchenmann, L., Curry, B., & Witek, M. A. G. (2023). Understanding the relationship between catchiness and groove: A qualitative study with popular music creators. *Music Perception*, 40(5), 353–372. <https://doi.org/10.1525/mp.2023.40.5.353>
- Burger, B., Thompson, M., Luck, G., Saarikallio, S., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, 4, Article 183. <https://doi.org/10.3389/fpsyg.2013.00183>
- Bowling, D. L., Graf Ancochea, P., Hove, M. J., & Fitch, W. T. (2019). Pupillometry of groove: Evidence for noradrenergic arousal in the link between music and movement. *Frontiers in Neuroscience*, 12, Article 1039. <https://doi.org/10.3389/fnins.2018.01039>
- Cho, E. (2016). Making reliability reliable: A systematic approach to reliability coefficients. *Organizational Research Methods*, 19(4), 651–682. <https://doi.org/10.1177/1094428116656239>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Danielsen, A., Nymoen, K., Langerød, M. T., Jacobsen, E., Johansson, M., & London, J. (2022). Sounds familiar(?): Expertise with specific musical genres modulates timing perception and micro-level synchronization to auditory stimuli. *Attention, Perception & Psychophysics*, 84(2), 599–615. <https://doi.org/10.3758/s13414-021-02393-z>
- Dotov, D., Bosnyak, D., & Trainor, L. J. (2021). Collective music listening: Movement energy is enhanced by groove and visual social cues. *The Quarterly Journal of Experimental Psychology: QJEP*, 74(6), 1037–1053. <https://doi.org/10.1177/1747021821991793>
- Dudenredaktion. (n.d.). eingrooven. In *Duden online*. Abgerufen am 1. April 2025 von <https://www.duden.de/rechtschreibung/eingrooven>
- Duman, D., Snape, N., Danso, A., Toiviainen, P., & Luck, G. (2024). Groove as a multidimensional participatory experience. *Psychology of Music*, 52(1), 93–116. <https://doi.org/10.1177/03057356231165327>
- Düvel, N., Labonde, P., Bechtold, T., Senn, O., & Kopiez, R. (2021). Experience of Groove Questionnaire: German translation and validation. *Music Perception*, 39(1), 83–99. <https://doi.org/10.1525/mp.2021.39.1.83>

- Düvel, N., Labonde, P., & Kopiez, R. (2022). Dissecting the groove experience: The role of instrumentation, low-level audio features, and personal musical background. *Psychomusicology: Music, Mind, and Brain*, 32(3–4), 109–122. <https://doi.org/10.1037/pmu0000296>
- Eden, A. (2020). Appreciation and eudaimonic reactions to media. In M.-L. Mares & E. Scharrer (Hrsg.), *The international encyclopedia of media psychology* (S. 1–9). John Wiley & Sons. <https://doi.org/10.1002/9781119011071.iemp0172>
- Etani, T., Miura, A., Kawase, S., Fujii, S., Keller, P. E., Vuust, P., & Kudo, K. (2024). A review of psychological and neuroscientific research on musical groove. *Neuroscience & Biobehavioral Reviews*, 158, Article 105522. <https://doi.org/10.1016/j.neubiorev.2023.105522>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Föderation Deutscher Psychologenvereinigungen. (2022). *Berufsethische Richtlinien des Berufsverbandes Deutscher Psychologinnen und Psychologen e. V. und der Deutschen Gesellschaft für Psychologie e. V.*
- Gabrielsson, A. (1973). Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology*, 14(1), 244–260. <https://doi.org/10.1111/j.1467-9450.1973.tb00118.x>
- Gonzalez-Sanchez, V. E., Zelechowska, A., & Jensenius, A. R. (2018). Correspondences between music and involuntary human micromotion during standstill. *Frontiers in Psychology*, 9, Article 1382. <https://doi.org/10.3389/fpsyg.2018.01382>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Hosken, F. (2020). The subjective, human experience of groove: A phenomenological investigation. *Psychology of Music*, 48(2), 182–198. <https://doi.org/10.1177/0305735618792440>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54–75. <https://doi.org/10.1037/a0024208>
- Joalla. (2024). *discogs\_client* [python Paket]. Github. [https://github.com/joalla/discogs\\_client](https://github.com/joalla/discogs_client)
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., . . . Johnson, A. R. (2022). *semTools: useful tools for structural equation modeling* (Version 0.5-6) [Software]. <https://CRAN.R-project.org/package=semTools>

- Juslin, P. N. (2019). Get into the groove: Rhythmic entrainment. In P. N. Juslin (Hrsg.), *Musical emotions explained: Unlocking the secrets of musical affect* (S. 275–286). Oxford University Press. <https://doi.org/10.1093/oso/9780198753421.003.0019>
- Juslin, P. N., & Sloboda, J. A. (Eds.). (2011). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press.
- Kern, H. J., & Lepa, S. (2014). Die “Techno-Schwelle”: Zur Bedeutsamkeit von Lautsprecherwiedergabe und Knochenleitung für die vestibulare Wirkung lauter Bassklänge im EDM-Clubkontext. In *Fortschritte der Akustik – DAGA 2014*. [http://pub.dega-akustik.de/DAGA\\_2014/data/articles/000021.pdf](http://pub.dega-akustik.de/DAGA_2014/data/articles/000021.pdf)
- Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, 2(3), 275–283. <https://doi.org/10.1525/can.1987.2.3.02a00010>
- Kowalewski, D. A., Kratzer, T. M., & Friedman, R. S. (2020). Social music: Investigating the link between personal liking and perceived groove. *Music Perception*, 37(4), 339–346. <https://doi.org/10.1525/mp.2020.37.4.339>
- Lepa, S. (2012). Was kann das Affordanzkonzept für eine Methodologie der Populärkulturforschung leisten? In M. S. Kleiner & M. Rappe (Hrsg.), *Methoden der Populärkulturforschung: Interdisziplinäre Perspektiven auf Film, Fernsehen, Musik, Internet und Computerspiele* (S. 273–298). Lit-Verlag.
- Levitin, D. J., Grahn, J. A., & London, J. (2018). The psychology of music: Rhythm and movement. *Annual Review of Psychology*, 69, 51–75. <https://doi.org/10.1146/annurev-psych-122216-011740>
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1(3), 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>
- Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, 24(2), 201–208. <https://doi.org/10.1525/mp.2006.24.2.201>
- Matthews, T. E., Stupacher, J., & Vuust, P. (2023). The pleasurable urge to move to music through the lens of learning progress. *Journal of Cognition*, 6(1), Article 55. <https://doi.org/10.5334/joc.320>
- Matthews, T. E., Witek, M. A. G., Heggli, O. A., Penhune, V. B., & Vuust, P. (2019). The sensation of groove is affected by the interaction of rhythmic and harmonic complexity. *PLoS ONE*, 14(1), Article e0204539. <https://doi.org/10.1371/journal.pone.0204539>
- Matthews, T. E., Witek, M. A. G., Lund, T., Vuust, P., & Penhune, V. B. (2020). The sensation of groove engages motor and reward networks. *NeuroImage*, 214, Article 116768. <https://doi.org/10.1016/j.neuroimage.2020.116768>
- Matyja, J. R., & Schiavio, A. (2013). Enactive music cognition: Background and research themes. *Constructivist Foundations*, 8(3), Article 3.
- McDonald, R. P. (1985). *Factor analysis and related methods*. Lawrence-Erlbaum.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.
- Merlini, M. (2020). A critical (and interdisciplinary) survey of popular music genre theories. *Interdisciplinary Studies in Musicology*, 20, 79–94. <https://doi.org/10.14746/ism.2020.20.7>

- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599–620.  
[https://doi.org/10.1207/S15328007SEM0904\\_8](https://doi.org/10.1207/S15328007SEM0904_8)
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS ONE*, 9(2), Article e89642. <https://doi.org/10.1371/journal.pone.0089642>
- O'Connell, S. R., Nave-Blodgett, J. E., Wilson, G. E., Hannon, E. E., & Snyder, J. S. (2022). Elements of musical and dance sophistication predict musical groove perception. *Frontiers in Psychology*, 13, Article 998321. <https://doi.org/10.3389/fpsyg.2022.998321>
- Pando-Naude, V., Matthews, T. E., Højlund, A., Jakobsen, S., Østergaard, K., Johnsen, E., Garza-Villarreal, E. A., Witek, M. A. G., Penhune, V., & Vuust, P. (2024). Dopamine dysregulation in Parkinson's disease flattens the pleasurable urge to move to musical rhythms. *The European Journal of Neuroscience*, 59(1), 101–118. <https://doi.org/10.1111/ejn.16128>
- Pfeleiderer, M. (2002). „It don't mean a thing if it ain't got that swing": Überlegungen zur mikrorhythmischen Gestaltung in populärer Musik. *Jahrbuch Musikpsychologie*, 16, 102–122.
- Pfeleiderer, M. (2006). *Rhythmus: Psychologische, theoretische und stilanalytische Aspekte populärer Musik*. transcript. <https://doi.org/10.1515/9783839405154>
- Pfeleiderer, M. (2010). Dimensionen der Groove-Erfahrung. Eine empirische Studie. *PopScriptum*, 11. <https://edoc.hu-berlin.de/handle/18452/21062>
- Raykov, T. (2012). Scale construction and development using structural equation modeling. In R. H. Hoyle (Hrsg.), *Handbook of structural equation modeling* (S. 472–492). Guilford Press.
- Roholt, T. C. (2014). *Groove. A phenomenology of rhythmic nuance*. Bloomsbury.
- Romkey, I. D., Matthews, T., Foster, N., Bella, S. D., & Penhune, V. B. (2025). The pleasurable urge to move to music is unchanged in people with musical anhedonia. *PLoS ONE*, 20(1), Article e0312030. <https://doi.org/10.1371/journal.pone.0312030>
- Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods*, 25(1), 6–14. <https://doi.org/10.1177/1094428120968614>
- Rose, D., Müllensiefen, D., Lovatt, P., & Orgs, G. (2022). The Goldsmiths Dance Sophistication Index (Gold-DSI): A psychometric tool to assess individual differences in dance experience. *Psychology of Aesthetics, Creativity, and the Arts*, 16(4), 733–745.  
<https://doi.org/10.1037/aca0000340>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Savalei, V. (2018). On the computation of the RMSEA and CFI from the mean-and-variance corrected test statistic with nonnormal data in SEM. *Multivariate Behavioral Research*, 53(3), 419–429. <https://doi.org/10.1080/00273171.2018.1455142>
- Schaal, N. K., Bauer, A.-K. R., & Müllensiefen, D. (2014). Der Gold-MSI: Replikation und Validierung eines Fragebogeninstrumentes zur Messung Musikalischer Erfahrungheit anhand einer deutschen Stichprobe. *Musicae Scientiae*, 18(4), 423–447.  
<https://doi.org/10.1177/1029864914541851>

- Schiavio, A., Witek, M. A. G., & Stupacher, J. (2024). Meaning-making and creativity in musical entrainment. *Frontiers in Psychology*, *14*, Article 1326773.  
<https://doi.org/10.3389/fpsyg.2023.1326773>
- Seeberg, A. B., Matthews, T. E., Højlund, A., Vuust, P., & Petersen, B. (2024a). *Beyond syncopation: The number of rhythmic layers shapes the pleasurable urge to move to music* [Preprint]. SSRN.  
<https://doi.org/10.2139/ssrn.4899990>
- Seeberg, A. B., Matthews, T. E., Højlund, A., Vuust, P., & Petersen, B. (2024b). *The pleasurable urge to move to music in cochlear implant users* [Preprint]. SSRN. <https://doi.org/10.2139/ssrn.4931906>
- Senn, O., Kilchenmann, L., Bechtold, T., & Hoesl, F. (2018). Groove in drum patterns as a function of both rhythmic properties and listeners' attitudes. *PLoS ONE*, *13*(6), Article e0199604.  
<https://doi.org/10.1371/journal.pone.0199604>
- Senn, O., Bechtold, T. A., Hoesl, F., & Kilchenmann, L. (2021). Taste and familiarity affect the experience of groove in popular music. *Musicae Scientiae*, *25*(1), 45–66.  
<https://doi.org/10.1177/1029864919839172>
- Senn, O., Bechtold, T., Rose, D., Câmara, G. S., Düvel, N., Jerjen, R., Kilchenmann, L., Hoesl, F., Baldassarre, A., & Alessandri, E. (2020). Experience of groove questionnaire: Instrument development and initial validation. *Music Perception*, *38*(1), 46–65.  
<https://doi.org/10.1525/mp.2020.38.1.46>
- Senn, O., Bechtold, T., Hoesl, F., Jerjen, R., Kilchenmann, L., Rose, D., Baldassarre, A., Sigrist, C., & Alessandri, E. (2023). An SEM approach to validating the psychological model of musical groove. *Journal of Experimental Psychology: Human Perception and Performance*, *49*(3), 290–305.  
<https://doi.org/10.1037/xhp0001087>
- Senn, O., Bechtold, T. A., Jerjen, R., Kilchenmann, L., & Hoesl, F. (2023). Three psychometric scales for groove research: Inner representation of temporal regularity, time-related interest, and energetic arousal. *Music & Science*, *6*. <https://doi.org/10.1177/20592043231185663>
- Senn, O., Hoesl, F., Bechtold, T. A., Kilchenmann, L., Jerjen, R., & Witek, M. (2024). Null effect of perceived drum pattern complexity on the experience of groove. *PLOS ONE*, *19*(11), Article e0311877. <https://doi.org/10.1371/journal.pone.0311877>
- Todd, N. P. M., & Cody, F. W. (2000). Vestibular responses to loud dance music: A physiological basis of the “rock and roll threshold”? *The Journal of the Acoustical Society of America*, *107*(1), 496–500. <https://doi.org/10.1121/1.428317>
- Windsor, W. L., & de Bézenac, C. (2012). Music and affordances. *Musicae Scientiae*, *16*(1), 102–120.  
<https://doi.org/10.1177/1029864911435734>
- Witek, M. A. G. (2017). Filling in: Syncopation, pleasure and distributed embodiment in groove. *Music Analysis*, *36*(1), 138–160. <https://doi.org/10.1111/musa.12082>
- Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*(4), Article e94446.  
<https://doi.org/10.1371/journal.pone.0094446>

## Anhang

**Tabelle 9**

*Oblimin-rotierte Ladungsmatrix der EFA bei der Entwicklung des DGA*

Itemtext	Faktor				1-Kommunalität
	1	2	3	4	
aggressiv	0.894				0.219
wild	0.872				0.166
gehetzt	0.828				0.330
weich	-0.806				0.326
heftig	0.801				0.328
energiegeladen	0.799				0.202
ruhelos	0.790			-0.320	0.350
beschleunigend	0.783				0.318
ruhig	-0.763				0.355
mechanisch	0.749				0.440
schnell	0.738				0.400
flackernd	0.732				0.432
hart	0.699		0.304		0.286
flexibel	-0.698	0.319			0.486
schaukelnd	-0.694				0.413
kraftvoll	0.690				0.323
träge	-0.668			-0.331	0.217
fließend	-0.658				0.539
aufregend	0.648				0.371
gehend	-0.597				0.592
zerfurcht	0.589		-0.450		0.483
hämmernd	0.576				0.499
markant	0.563				0.594
stotternd	0.559	-0.536			0.441
intensiv	0.544		0.302		0.509
zunehmend	0.534	-0.351		0.387	0.481
klopfend	0.524				0.699
kompliziert	0.509		-0.463		0.560
schwingend	-0.496				0.708
elastisch	-0.489	0.464			0.598
fest	0.483				0.551
angespannt	0.474				0.751
ereignisreich	0.458		0.322		0.579
saugend	-0.457				0.680
glatt	-0.440		0.372		0.608
kunstvoll	0.432				0.764
vorwärtstreibend	0.417		0.382		0.553
zusammengesetzt					0.901
pulsierend					0.807
hinkend		-0.717			0.492
lang		-0.685			0.452
bewegungsanregend		0.659			0.467
beweglich		0.603			0.616
statisch		-0.531			0.729
zögerlich	-0.417	-0.513			0.412
kontrastreich		0.502			0.630
monoton		-0.498	0.366		0.653
lebendig		0.489			0.454

Itemtext	Faktor				1-Kommunalität
	1	2	3	4	
stoppend		-0.471			0.728
vital		0.451		0.449	0.400
springend	0.375	0.441			0.512
abwechslungsreich		0.392	-0.327		0.747
zurückschnellend					0.961
regelmässig			0.864		0.244
einheitlich		-0.385	0.685		0.447
klar			0.674		0.475
geordnet			0.651		0.552
schwungvoll			0.632	0.362	0.416
uneinheitlich			-0.625	-0.369	0.449
beherrscht			0.586	-0.343	0.537
diffus			-0.537		0.683
beständig			0.515		0.690
akzentuiert	0.393		0.485	0.331	0.317
einfach		-0.426	0.468		0.623
bestimmt	0.372		0.429		0.477
rhythmisch		0.332	0.387		0.668
graziös	0.312		-0.326		0.805
schwer				-0.861	0.275
freudig				0.693	0.344
verspielt				0.625	0.561
feierlich		-0.336		0.609	0.556
fröhlich		0.317	0.316	0.555	0.354
frei		-0.466		0.523	0.537
stampfend	0.461			-0.483	0.594
begeistert		0.429		0.480	0.316
kriechend				-0.385	0.525
hüpfend				0.310	0.690

Anmerkung. Ladungen < .3 unterdrückt und nach Größe sortiert.

**Tabelle 10**

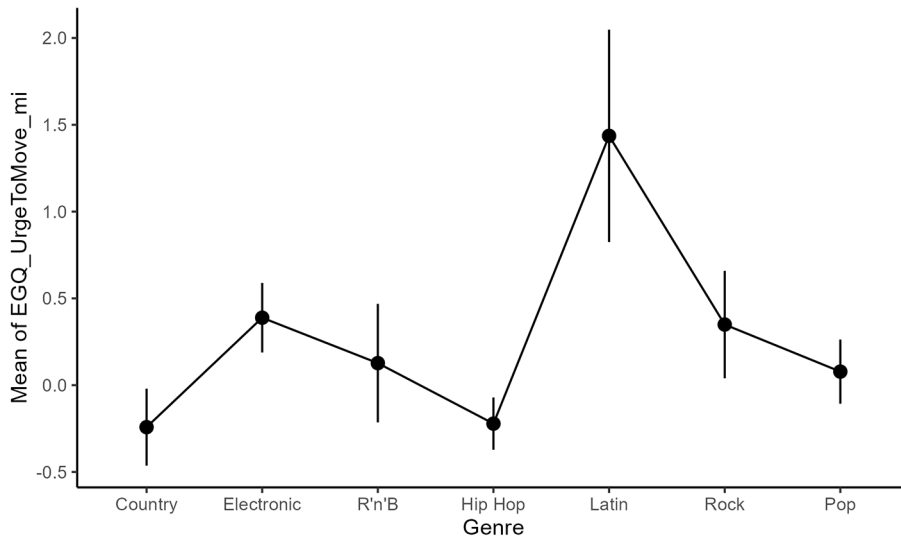
Geschätzte LMM-Effektkontraste für EGQ Urge to move (Mittelwertindizes): Genrebeurteilungen gegen Gesamtmittelwert

Songgenre	Vergleichswert	Differenz	SE	95% CI	t (df = 1760)	p
Country	MW (alle Genres inkl. Pop)	-0.52	0.1	[-0.71, -0.32]	-5.08	< .001
Electronic	MW (alle Genres inkl. Pop)	0.11	0.09	[-0.07, 0.30]	1.24	.216
R'n'B	MW (alle Genres inkl. Pop)	-0.15	0.15	[-0.44, 0.15]	-0.97	.332
HipHop	MW (alle Genres inkl. Pop)	-0.5	0.07	[-0.64, -0.35]	-6.74	< .001
Latin	MW (alle Genres inkl. Pop)	1.16	0.27	[0.64, 1.68]	4.37	< .001
Rock	MW (alle Genres inkl. Pop)	0.08	0.14	[-0.20, 0.35]	0.55	.585

Anmerkung. Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten.

**Abbildung 9**

Geschätzte Genremittelwerte des LMM zu EGQ Urge to move (Mittelwertindizes)



Anmerkung. Fehlerbalken repräsentieren 95% Konfidenzintervalle.

**Tabelle 11**

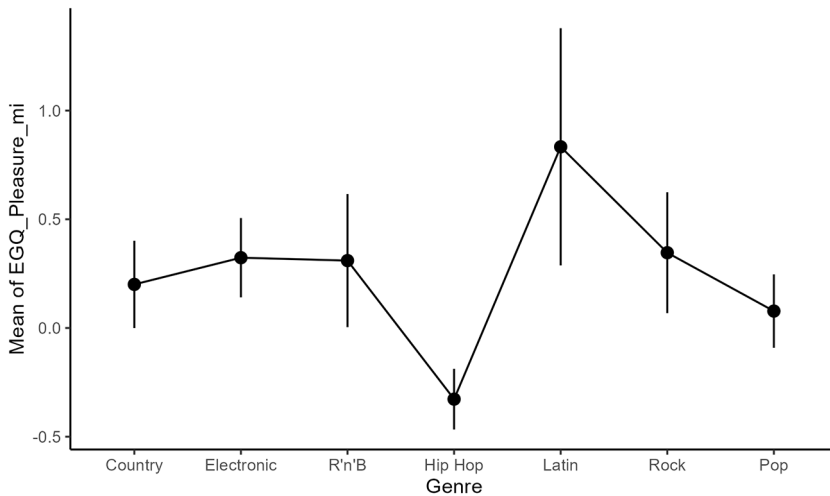
Geschätzte LMM-Effektkontraste für EGQ Pleasure (Mittelwertindizes): Genrebeurteilungen gegen Gesamtmittelwert

Songgenre	Vergleichswert	Differenz	SE	95% CI	<i>t</i> ( <i>df</i> = 1760)	<i>p</i>
Country	MW (alle Genres inkl. Pop)	-0.05	0.09	[-0.23, 0.13]	-0.57	.568
Electronic	MW (alle Genres inkl. Pop)	0.07	0.08	[-0.09, 0.23]	0.86	.388
R'n'B	MW (alle Genres inkl. Pop)	0.06	0.13	[-0.21, 0.32]	0.43	.666
HipHop	MW (alle Genres inkl. Pop)	-0.58	0.07	[-0.71, -0.45]	-8.87	< .001
Latin	MW (alle Genres inkl. Pop)	0.58	0.24	[0.12, 1.05]	2.46	.014
Rock	MW (alle Genres inkl. Pop)	0.09	0.12	[-0.15, 0.34]	0.76	.445

Anmerkung. Sämtliche anderen Modellkoeffizienten wurden zur Kontrastschätzung auf dem Mittelwert gehalten.

**Abbildung 10**

Geschätzte Genremittelwerte des LMM zu EGQ Pleasure (Mittelwertindizes)



Anmerkung: Fehlerbalken repräsentieren 95% Konfidenzintervalle.



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



Leibniz-Institut für  
Psychologie

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