

A Systematic Literature Review on Empirical Applications of Gordon's Music Learning Theory and Aptitude Measures: Sino-International Perspectives (2005–2025)

Yutong Gu, Sajastanah Bin Imam Koning

Faculty of Music and Performing Arts, Sultan Idris Education University, Malaysia

Corresponding Author Email: sajastanah@fmstp.upsi.edu.my

To Link this Article: <http://dx.doi.org/10.6007/IJARPED/v14-i2/25342> DOI:10.6007/IJARPED/v14-i2/25342

Published Online: 12 May 2025

Abstract

This study systematically reviews empirical research on Edwin Gordon's Music Learning Theory (MLT) and associated music aptitude measures conducted between 2005 and 2025, particularly emphasizing contrasting international and Chinese research contexts. Gordon's MLT, centered on audiation, has significantly influenced music pedagogy and aptitude assessment globally. However, the extent and rigor of empirical applications differ widely across cultural and educational systems. This review seeks to consolidate current evidence and illuminate how MLT's utility has been examined, implemented, and interpreted internationally and in China. Following PRISMA guidelines, a comprehensive search across Scopus, ERIC, and CNKI yielded 38 empirical studies meeting inclusion criteria. Thematic synthesis revealed three dominant trends: (1) MLT-based interventions yield consistently positive outcomes in musical, cognitive, and affective domains; (2) international studies employ advanced psychometric analyses, while Chinese research remains largely in exploratory or theoretical stages; and (3) audiation aptitude shows strong interdisciplinary correlations, though such links are underexplored in China. This review highlights an urgent need to localize and empirically validate Gordon's frameworks within Chinese education. It contributes to bridging the gap between theory and practice and offers implications for curriculum reform, teacher training, and cross-cultural educational research. The study is particularly beneficial for music educators, researchers, policymakers, and curriculum developers seeking evidence-based approaches to advance music aptitude development in diverse educational contexts.

Keywords: Music Learning Theory, Edwin Gordon, Audiation, Music Aptitude, Empirical Research, Cross-Cultural Comparison

Introduction

Music aptitude development is a fundamental component of comprehensive music education. In this regard, Edwin Gordon's Music Learning Theory (MLT) has emerged as one of the most influential frameworks in understanding how musical understanding is acquired,

particularly through audiation—the cognitive capacity to think musically without external sound (Gordon, 2012). Over the past two decades, Gordon’s theory and his suite of standardized aptitude assessments have provided both a theoretical foundation and practical tools for nurturing musical potential in educational settings, comprising the Primary Measures of Music Audiation (PMMA), Intermediate Measures of Music Audiation (IMMA), Advanced Measures of Music Audiation (AMMA), and the Musical Aptitude Profile (MAP). These evaluations have gained extensive use in educational practice and empirical study in North America and Europe.

The importance of this topic lies in its far-reaching implications for both music pedagogy and interdisciplinary learning. Empirical applications of MLT have shown significant correlations between audiation-based instruction and improvements in rhythm accuracy, tonal discrimination, language development, mathematical reasoning, and socio-emotional learning (Bergee, 2005; Gromko, 2004; Ilari, 2016). These findings reinforce the utility of MLT not only in developing musical skill, but also in supporting broader cognitive and educational development—thus making it a strategically valuable research domain for educators, psychologists, and curriculum designers alike.

Globally, empirical research based on MLT has explored diverse educational issues, including the predictive significance of musical aptitude on learning outcomes and the creation of tailored instructional strategies in early childhood and primary education (Bergee, 2005; Gromko, 2004; Ilari, 2016). MLT has been implemented in the United States, Canada, Brazil, and certain regions of Europe via experimental designs, longitudinal studies, and classroom-based treatments. These empirical endeavors not only corroborate the theoretical assertions of MLT but also provide insights into how audiation-based sequencing can improve rhythmic and tonal development in learners. Conversely, China’s engagement with Gordon’s work has remained largely theoretical, with most efforts concentrated on literature interpretation, curriculum proposals, and limited pilot projects. Empirical research validating MLT’s tools or outcomes in Chinese classrooms remains scarce. The reception and implementation of Gordon’s theory inside the Chinese academic sphere have adhered to a distinctly divergent path. Since its introduction to Chinese readers in the early 1980s via translated works and theoretical expositions (e.g., Lin, 1981), the majority of research has concentrated on theoretical interpretation, conceptual localization, and curricular integration.

This gap is significant. Without empirical validation, the potential of Gordon’s theory to transform music education in China is limited. The absence of large-scale studies prevents educators from understanding how MLT might align with or challenge existing pedagogical traditions. As a result, policymakers and practitioners lack the data needed to implement audiation-based approaches confidently and systematically.

Furthermore, this divergence mirrors a wider pattern in international education research: the selective adaptation of Western pedagogical frameworks without thorough contextual testing. While Western studies offer robust data-driven insights, Chinese literature on MLT is predominantly conceptual or speculative. This raises critical questions about how global theories are received and operationalized in local educational systems.

Accordingly, this systematic review addresses the need to: Consolidate empirical findings from both international and Chinese sources on MLT and Gordon's aptitude tools; Identify methodological trends, limitations, and innovations in how these studies are conducted; Illuminate pathways for developing culturally responsive and psychometrically robust music education practices in China and beyond.

To guide this investigation, the following research questions are proposed:

- i. What empirical trends define the global and Chinese implementations of Gordon's Music Learning Theory and aptitude assessments?
- ii. How do methodological practices and evaluation tools differ between international and Chinese research contexts?
- iii. What research gaps exist in China regarding MLT, and how do they compare with global progress?

By exploring these questions, this study aims to enhance the utility, visibility, and application of Gordon's theory in culturally diverse settings. The findings will provide valuable insights for educators, researchers, and policymakers seeking to design more inclusive, data-driven, and pedagogically effective music curricula. Ultimately, the study seeks to support the global advancement of music learning grounded in empirical evidence and contextual relevance.

Literature Review

Edwin Gordon's Music Learning Theory (MLT) has profoundly impacted global music education. At the core of MLT is the notion of audiation, which is the cognitive mechanism by which individuals internally perceive and understand music even when it is not present (Gordon, 2012). Gordon created organized sequences and systematic evaluations, including the Primary Measures of Music Audiation (PMMA), Intermediate Measures of Music Audiation (IMMA), Advanced Measures of Music Audiation (AMMA), and Music Aptitude Profile (MAP), to measure musical potential and guide instructional methodologies (Hanson, 2019).

MLT has been used globally across diverse age demographics and educational environments, showcasing its effectiveness in evaluating musical aptitude, enhancing teaching methodologies, and facilitating musical skill advancement (Di Mascio et al., 2020; Çenberci & Tufan, 2023). Nonetheless, worldwide empirical studies underscore persistent issues about the validity and applicability of Gordon's measures, especially concerning their ability to exclusively assess audiation as opposed to intersecting cognitive domains like tone memory (Platz et al., 2022).

Empirical investigations utilizing Gordon's MLT reveal variations in study emphasis and methodological rigor between China and worldwide environments. Global studies thoroughly investigate the psychometric validity of Gordon's assessments (Degé et al., 2017; Platz et al., 2022), encompassing their correlations with non-musical cognitive areas, including language, mathematics, and social-emotional growth (McDonel, 2015; Çenberci & Tufan, 2023). In contrast, Chinese empirical research mostly focuses on localized adaptations and preliminary investigations of MLT's educational applications rather than comprehensive psychometric validation (Zhou, 2021; Luo, 2022).

Recent study in China demonstrates efforts to localize MLT in early childhood and primary education environments. Research has utilized experimental methodologies to evaluate the efficacy of audiation-based education in enhancing musical talents and associated cognitive skills, including focus and memory (Lu, 2023; Deng, 2022; Zhang, 2022). Nonetheless, these studies frequently remain preliminary, characterized by smaller sample sizes and shorter periods in comparison to worldwide research. Furthermore, international research sometimes uses sophisticated statistical techniques, like Rasch modeling, hierarchical regression, and Bayesian analyses, to thoroughly examine the validity and reliability of Gordon's measures (Platz et al., 2022; Investigating Musical Aptitude Examination, ERIC). Chinese studies, however, have demonstrated minimal involvement with advanced psychometric procedures, underscoring a notable deficiency in methodological sophistication (Lv, 2022; Li, 2018).

In conclusion, although Gordon's MLT and aptitude assessments are widely acknowledged and well validated globally, their practical implementations in Chinese contexts are still in the early stages. The literature highlights an urgent necessity for more comprehensive and extensive validation research in China to ascertain the cross-cultural validity and effectiveness of Gordon's educational methods and theories.

Methodology

Identification

The literature search was conducted across three academic databases representing both international and Chinese scholarly output: Scopus, ERIC, and CNKI. These platforms were selected for their comprehensive indexing of peer-reviewed journal articles in education and music pedagogy.

To ensure a wide yet focused coverage of relevant studies, a Boolean search strategy was designed around three conceptual domains:

- **Theory and theorist:** "Edwin Gordon" OR "Music Learning Theory" OR "MLT"
- **Measurement tools and core concepts:** "music aptitude" OR "music aptitude test" OR "PMMA" OR "IMMA" OR "AMMA" OR "MAP" OR "audiation"

In Scopus, Boolean logic was applied to search titles, abstracts, and keywords using combinations of the above terms. In CNKI, the equivalent Chinese terms “埃德温·戈登” were employed to identify literature written in Chinese. Given that empirical research in China is still in its infancy, and in order to comprehensively reflect research trends in the Chinese context, this study includes selected master's and doctoral theses that have undergone quality review and demonstrate strong empirical rigor, clear methodology, and standardized structure. Due to the limited result set in ERIC when using full Boolean combinations, an iterative keyword-specific search strategy was adopted. Seven key terms—"Edwin Gordon", "Music learning theory (MLT)", "audiation", "music aptitude", "PMMA", "IMMA", and "AMMA"—were searched separately. The individual result sets were merged and processed using Zotero to remove duplicates. This approach ensured a broader and more accurate inclusion of relevant records, especially given variation in keyword usage and indexing practices within ERIC.

As shown in Table 1, the full search strings and strategies for each database are presented. 142 records were retrieved from Scopus, 240 from ERIC (after deduplication), and

66 from CNKI, yielding an initial corpus of 448 articles for further screening. All citations were imported into Zotero for reference management and subsequent eligibility evaluation.

Table 1

Identification of keywords and the exploration of associated terms in databases

Scopus	(TITLE-ABS-KEY ("Edwin Gordon" OR "Gordon Music Learning Theory" OR "Music Learning Theory" OR "MLT") AND TITLE-ABS-KEY ("music aptitude" OR "music aptitude test" OR "IMMA" OR "PMMA" OR "AMMA" OR "MAP" OR "audiation")) AND PUBYEAR > 2004 AND PUBYEAR < 2026
ERIC	Individual searches using: "Edwin Gordon", "audiation", "PMMA", "IMMA", "AMMA", and "music aptitude"; deduplicated total = 240
CNKI	(TITLE-ABS-KEY: "埃德温戈登")

Screening

In the screening phase, each of the 448 records was independently assessed based on its title and abstract. The goal was to exclude studies that were irrelevant or did not meet the review's inclusion criteria. All retrieved records were imported into Zotero for duplicate detection during the initial screening phase. After inspection, a small number of overlapping entries were found and excluded. The screening process proceeded in two distinct rounds, designed to refine the initial pool of literature and ensure relevance to the study's objectives.

In the first round, a total of 294 articles were excluded based on their titles and abstracts, primarily due to disciplinary irrelevance (e.g., unrelated uses of the acronym "MLT" in fields such as linguistics, medicine, or geology). The remaining 154 articles were retained for further review.

In the second round, these 154 articles were assessed more stringently. Inclusion and exclusion criteria were applied to evaluate journal articles published between 2005 and 2025, and those written in either English or Chinese were retained. Publications in the form of reviews, meta-analyses, book chapters, theses, and conference proceedings were excluded. In the case of Chinese-language literature retrieved from CNKI, master's and doctoral theses and journals were retained during screening only when they satisfied the basic formal criteria and demonstrated potential relevance. This two-step screening process resulted in 126 studies being retained for full-text examination during the eligibility phase. The selection criteria applied throughout this stage are summarized in Table 2.

Table 2

The selection search criteria

Criterion	Included	Excluded
Language	English, Chinese	Other languages
Publication Year	2005–2025	Before 2005 or beyond 2025
Literature Type	Peer-reviewed journal articles; Chinese master's/doctoral theses (empirical only)	Theoretical papers, reviews, meta-analyses, books, book chapters, dissertations (English), conference proceedings
Disciplinary Scope	Music education, music psychology, educational research	Medicine, geology, linguistics, and non-music-related domains
Publication Status	Final published articles and archived academic theses (Chinese)	In press, early access, unpublished informal content

Eligibility

During the third stage, known as the eligibility phase, a total of 125 items was assembled. A thorough assessment was performed on the titles and primary content of all publications to determine their compliance with the inclusion criteria and their relevance to the research aims of this study. As a result, 87 journals were removed from the study due to conflicting inclusion criteria. The complete text has been excluded from this analysis due to the presence of non-empirical research or irrelevant Gordon ($n=75$). Furthermore, the title does not demonstrate a substantial correlation ($n=12$), as it lacks empirical evidence.

While Scopus and ERIC yielded exclusively peer-reviewed journal articles, the CNKI database presented a unique challenge. Systematic search revealed a limited presence of peer-reviewed empirical journal articles in the Chinese context (only three studies), yet a growing body of relevant master's theses ($N = 10$) that demonstrated clear empirical design and analytical rigor. Therefore, to ensure comprehensiveness and regional inclusivity, this inclusion strategy addressed the contextual research gap in Chinese music education and ensured meaningful representation of localized scholarly perspectives. After this final eligibility step, as shown in Figure 1, 38 studies were selected for synthesis.

Data Abstraction and Analysis

This study applied an integrative analytical approach to combine findings across several research types, including quantitative, qualitative, and mixed-method investigations. The final pool of 44 eligible studies was systematically reviewed to extract data relevant to Gordon's Music Learning Theory (MLT) and his other related theories. Data abstraction focused on recording publishing details, study setting, participant demographics, educational levels, methodological design, and the precise use of Gordon-related frameworks or instruments. Each study was classified manually to find recurring features such as tone and rhythmic audiation, the use of standardized testing, instructional interventions, and cognitive or educational outcomes.

Thematic analysis was undertaken to categorize research according to educational level, type of aptitude tool utilized, and geographical origin. Comparative interpretation was utilized to highlight disparities between international and Chinese empirical research, particularly in methodological depth, area of application, and pedagogical emphasis. All findings were

independently checked and validated to guarantee analytical consistency and thematic clarity, giving a systematic basis for the synthesis offered in the discussion section.

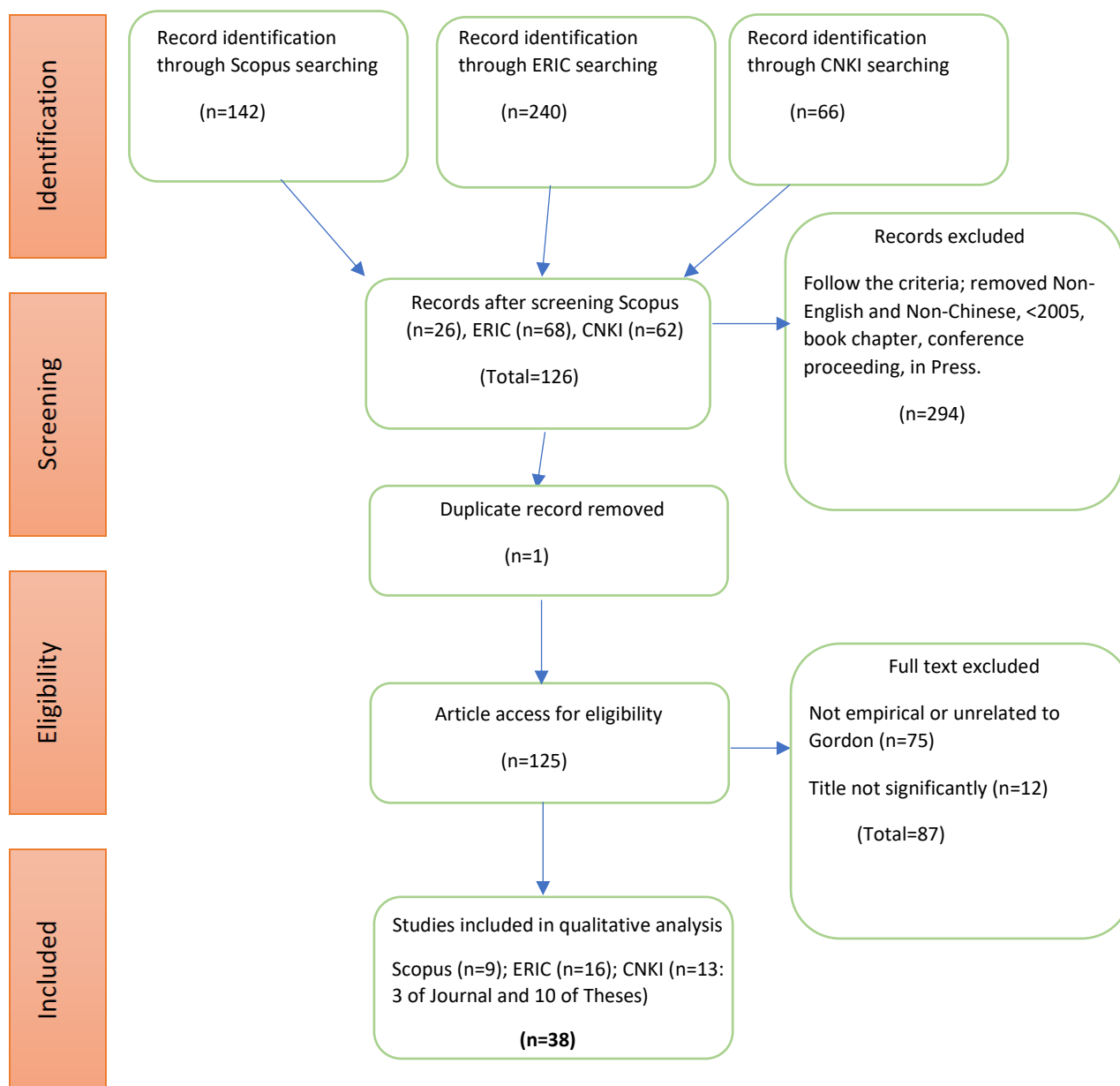


Figure 1. Flow diagram of the proposed searching study

Results

Theme 1: Educational Outcomes of MLT-based Pedagogical Interventions

A prominent feature discerned from the analyzed empirical investigations is the educational efficacy of pedagogical interventions grounded in Gordon's Music Learning Theory (MLT). Both international and Chinese research consistently demonstrates substantial positive effects across various educational environments and age demographics. Researchers globally have extensively applied MLT-based curriculum in diverse educational contexts, showcasing significant enhancements in musical skills. Research from Turkey, Italy, and Greece demonstrates that MLT-guided instruction markedly improves students' tonal, rhythmic, and composite musical abilities. Moreover, these interventions have exhibited supplementary advantages beyond musical abilities, enhancing children's social-emotional skills, self-

assurance, and cooperative behaviors in classroom environments (Çenberci & Tufan, 2023; Di Mascio et al., 2020; Esmergül & Önder, ERIC).

Concurrent results from Chinese research corroborate this viewpoint. Chinese academics have tailored MLT concepts to accommodate local educational requirements, resulting in significant achievements in early childhood and primary education, albeit generally being smaller in scale. Lu (2023) implemented audiation-centered therapies in kindergartens, noting substantial improvements in young children's musical perception, concentration, memory, and expressive skills. Luo (2022) examined the viability of integrating Gordon's theory into fundamental music courses, recording notable enhancements in students' musical aptitude, as demonstrated by pre- and post-test comparisons utilizing MAP. Wang's (2022) research demonstrates that structured MLT-informed teaching practices markedly improved students' capacity to understand rhythm and melodic patterns in regional elementary school settings.

These collective findings highlight the versatility and efficacy of MLT-based instructional tactics in various educational contexts. Nonetheless, the existing corpus of Chinese research, while promising, predominantly emphasizes short-term interventions with limited sample sizes. Consequently, additional longitudinal research involving bigger, diverse populations are essential to validate and generalize these educational results throughout the wider Chinese educational context.

Theme 2: Psychometric Validation and Cultural Adaptation of Music Aptitude Assessments

The second significant theme is the psychometric characteristics and cross-cultural applicability of Gordon's music aptitude assessments. The global literature underscores the necessity of comprehensive psychometric evaluations, meticulously assessing Gordon's assessments, including AMMA, IMMA, and PMMA, utilizing robust statistical techniques such as Rasch modeling, hierarchical regression, and factor analysis. These investigations often uncover significant concerns related to construct validity, measurement invariance, and cultural equity. Platz et al. (2022) meticulously evaluated AMMA's discriminant validity, indicating that it assesses both audiation and tonal memory, hence questioning its intended conceptual integrity. Degé et al. (2017) similarly identified modest correlations between IMMA and AMMA scores in children aged 9 to 13, indicating possible age-related psychometric constraints and prompting inquiries regarding the measures' validity across developmental phases.

In the Chinese setting, psychometric validation research is still comparatively underdeveloped. The majority of Chinese research undertaken to date have predominantly concentrated on preliminary translations, modifications, and fundamental reliability assessments. Zhou (2021) and Lyu (2022) investigated the localized applicability of MAP and PMMA in primary and kindergarten educational environments, respectively, offering initial evidence for the practicality of these assessments in Chinese classrooms. Nevertheless, these investigations infrequently encompass thorough psychometric analyses that would rigorously validate the test constructs, such as performing confirmatory factor analyses or investigating differential item functioning across various cultural groups.

The significant disparity between foreign and Chinese psychometric validation studies highlights a crucial subject for further investigation. Future studies must systematically examine cultural equivalence, validity across various age demographics, and sensitivity to local cultural nuances to enable the meaningful use and appropriate interpretation of Gordon's measurements within Chinese educational contexts. A comprehensive psychometric study, utilizing sophisticated statistical modeling, will improve the reliability and validity of music aptitude evaluations, facilitating their effective incorporation into China's educational assessment frameworks.

Theme 3: Interdisciplinary Cognitive Associations with Musical Proficiency

The third significant aspect discovered is the strong association between musical talent, as evaluated by Gordon's measurements, and various cognitive domains beyond music. Empirical research from both international and Chinese contexts continuously underscores the significance of audiation as a vital integrative cognitive function, impacting several educational outcomes, including language, mathematics, and overall academic achievement.

Research worldwide has carefully examined the correlation between audiation and broader cognitive areas, highlighting its connection to language competencies, including phonological awareness, pronunciation precision in second-language learners, and early literacy abilities (Hanson, 2019; ERIC, 2023). McDonel (2015) presented compelling evidence connecting early children's rhythmic audiation skills with developing mathematical cognition, indicating common cognitive processing mechanisms for both music and mathematics. These interdisciplinary research studies demonstrate audiation's ability to enhance cognitive flexibility, working memory, and abstract reasoning abilities.

Correspondingly, nascent Chinese research corresponds with global discoveries, albeit presently limited in scope and profundity. Initial Chinese research has commenced investigating the correlations between musical aptitude and broader cognitive abilities, including language acquisition, overall learning behaviors, and academic performance. Li (2018) identified a notable association between the musical aptitude of senior kindergarten students, as assessed by the PMMA, and their overall learning characteristics, encompassing persistence, attentiveness, and cognitive engagement. Tian (2021) and Wang (2021) investigated the effects of audiation-focused teaching approaches on junior high school students' music skills and academic performance, indicating enhancements in students' classroom involvement, concentration, and general learning attitudes.

The interdisciplinary cognitive connections emphasize audiation's extensive educational significance, transcending just musical ability development to encompass overall cognitive improvement. Nonetheless, the majority of Chinese research are still exploratory and preliminary, with smaller sample numbers and less extensive cognitive evaluations than their foreign equivalents. Future studies in China should carefully examine these interdisciplinary connections through extensive, longitudinal studies utilizing advanced cognitive evaluations. By thoroughly examining the role of audiation in educational outcomes, researchers may offer robust empirical evidence for the incorporation of music education and cognitive skill development methodologies into the general education curriculum in China.

Conclusions

This systematic review aimed to examine and consolidate empirical findings from 2005 to 2025 about Edwin Gordon's Music Learning Theory (MLT) and its related music aptitude assessments, contrasting research undertaken in China with that in the international arena. This study, employing a stringent PRISMA framework for selection and thematic analysis of empirical papers sourced from Scopus, ERIC, and CNKI databases, produced noteworthy findings that elucidate the evolution, implementation, and persistent issues related to MLT-based research.

The evidence strongly indicates that pedagogical interventions based on MLT have consistently yielded favorable educational results in both international and Chinese contexts. International research has thoroughly confirmed the effectiveness of MLT-based interventions in improving musical ability and related cognitive areas, including language learning, mathematical reasoning, and social-emotional development. Chinese research, while nascent and less comprehensive, similarly indicate promising advancements, especially in early childhood and elementary education contexts. Chinese research is predominantly marked by limited sample sizes and brief intervention durations, highlighting a significant necessity for extensive longitudinal investigations to thoroughly corroborate these initial findings.

Secondly, psychometric assessments of Gordon's aptitude metrics indicate significant disparities between foreign and Chinese studies. International research often use sophisticated statistical modeling to examine construct validity, discriminant validity, and cross-cultural adaptability of instruments like AMMA, IMMA, and PMMA. These investigations reveal persistent psychometric issues, notably the confounding of audiation with other cognitive variables such as tone memory. Conversely, Chinese research initiatives to adapt and validate these measures are still in the nascent stages, frequently confined to fundamental reliability evaluations without comprehensive construct validation. This gap indicates an urgent necessity for Chinese scholars to adopt more advanced psychometric approaches to guarantee the cultural relevance and scientific integrity of Gordon's assessments in local educational contexts.

The review emphasizes the significant and frequently undervalued function of musical aptitude—particularly audiation—as a complex cognitive ability with extensive ramifications beyond music instruction. International and rising Chinese studies both highlight substantial links between musical aptitude and cognitive abilities, including phonological awareness, overall academic capabilities, and socio-emotional competencies. These findings support a more expansive understanding of music education, viewing it not alone as an aesthetic or technical endeavor, but as an essential factor in comprehensive cognitive and educational growth.

The findings of this assessment collectively illustrate a diverse but uneven worldwide research landscape. Globally, Gordon's MLT and aptitude assessments have prompted vast, methodologically sound empirical research, whereas in China, empirical investigation is still in a nascent phase but exhibits significant potential. To bridge this gap, Chinese academics must conduct larger, more methodologically rigorous investigations that incorporate cross-cultural psychometric validation alongside longitudinal educational interventions.

This review confirms the lasting significance and pedagogical merit of Gordon's Music Learning Theory. Future research should focus on enhancing the incorporation of MLT into global educational systems, namely by creating culturally responsive assessment instruments and pedagogical methods that honor the linguistic, cognitive, and musical variety of students. By doing so, the global music education community may fulfill Gordon's vision of fostering audiation as a universal human ability, thus enhancing the intellectual, emotional, and social experiences of students across diverse cultural contexts.

Acknowledge

We would like to express my sincere appreciation to the University Pendidikan Sultan Idris (UPSI) for granting me the opportunity to undertake a doctoral program in music education. By granting me access to the necessary resources for this paper, the university provided me with enormous support.

References

- Brodsky, W., Kessler, Y., Rubinstein, B. S., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: notational audiation. *Journal of Experimental Psychology: Human Perception and Performance*, 34(2), 427.
- Brophy, T. S. (2005). A longitudinal study of selected characteristics of children's melodic improvisations. *Journal of Research in Music Education*, 53(2), 120-133.
- Burton, S. L. (2017). Making music mine: the development of rhythmic literacy. *Music Education Research*, 19(2), 133-142.
- Çenberci, S., & Tufan, E. (2023). Effect of music education based on Edwin E. Gordon's theory on children's developmental music aptitude and social emotional learning skills. *Music Education Research*, Advance online publication. <https://doi.org/10.1080/14613808.2023.2202360>
- Culp, M. E. (2017). The relationship between phonological awareness and music aptitude. *Journal of Research in Music Education*, 65(3), 328-346.
- Dégé, F., Patscheke, H., & Schwarzer, G. (2017). Associations between two measures of music aptitude: Are the IMMA and the AMMA significantly correlated in a sample of 9- to 13-year-old children? *Musicae Scientiae*, 21(4), 465-478. <https://doi.org/10.1177/1029864916670205>
- Esmergül, P., & Önder, V. B. (2022). Adaptation of Edwin Gordon's Primary and Intermediate Measures of Music Audiation Test to Turkish and Validity Reliability Study. *Turkish Online Journal of Educational Technology-TOJET*, 21(1), 18-34.
- Gübeş, N. Ö. (2021). Investigating Musical Aptitude Examination with a Many-Facet Rasch Model. *Participatory Educational Research*, 8(4), 334-345.
- Hanson, J. (2019). Meta-analytic evidence of the criterion validity of Gordon's music aptitude tests in published music education research. *Journal of Research in Music Education*, 67(2), 193-213. <https://doi.org/10.1177/0022429419838386>
- Hornbach, C. M., & Taggart, C. C. (2005). The relationship between developmental tonal aptitude and singing achievement among kindergarten, first-, second-, and third-grade students. *Journal of Research in Music Education*, 53(4), 322-331.
- Lowe, G., & Belcher, S. (2012). Direct instruction and music literacy: one approach to augmenting the diminishing?. *Australian Journal of Music Education*, (1), 3-13.
- Lu, Y. T. (2023). *Application and study of Edwin Gordon's audiation theory in music education for children aged 4-6* [Master's thesis, East China Normal University]. CNKI.

- Luo, P. N. (2022). *Research on Gordon's MAP test and feasibility analysis of its theory integration into basic music courses* [Master's thesis, Southwest University]. CNKI.
- Maerker Garner, A. (2009). Singing and moving: Teaching strategies for audiation in children. *Music educators journal*, 95(4), 46-50.
- McCormack, B. A., Klopper, C., Kitson, L., & Westerveld, M. (2018). The potential for music to develop pronunciation in students with English as an Additional Language or Dialect (EAL/D). *Australian Journal of Music Education*, 52(1), 43-50.
- McDonel, J. S. (2015). Exploring learning connections between music and mathematics in early childhood. *Early Childhood Research & Practice*, 17(1), Article 8.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372.
- Pan, S. S. (2019). Development and insights of 'Playing Music' classrooms in kindergartens. *Preschool Education Research*. Retrieved from CNKI.
- Pan, S. S. (2020). Practical insights into the localization of Edwin Gordon's early childhood music learning theory in China. *Early Childhood Education Research*. Retrieved from CNKI.
- Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2022). Measuring audiation or tonal memory? Evaluation of the discriminant validity of Edwin E. Gordon's "Advanced Measures of Music Audiation." *Psychology of Music*, Advance online publication. <https://doi.org/10.1177/03057356221086440>
- Pollatou, E., Karadimou, K., & Gerodimos, V. (2005). Gender differences in musical aptitude, rhythmic ability and motor performance in preschool children. *Early child development and care*, 175(4), 361-369.
- Reifinger Jr, J. L. (2018). The relationship of pitch sight-singing skills with tonal discrimination, language reading skills, and academic ability in children. *Journal of Research in Music Education*, 66(1), 71-91.
- Ruthsatz, J., Detterman, D., Griscom, W. S., & Cirullo, B. A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence*, 36(4), 330-338.
- Sarkis-Onofre, R., Catalá-López, F., Aromataris, E., & Lockwood, C. (2021). How to properly use the PRISMA Statement. *Systematic Reviews*, 10, 1-3.
- Selçuk, A. A. (2019). A guide for systematic reviews: PRISMA. *Turkish archives of otorhinolaryngology*, 57(1), 57.
- Stamou, L., Schmidt, C. P., & Humphreys, J. T. (2010). Standardization of the gordon primary measures of music audiation in greece. *Journal of Research in Music Education*, 58(1), 75-89.
- Verdis, A., & Sotiriou, C. (2018). The psychometric characteristics of the Advanced Measures of Music Audiation in a region with strong non-Western music tradition. *International Journal of Music Education*, 36(1), 69-84.
- Wang, S. N. (2022). *Classroom practice research on Edwin Gordon's Music Learning Theory in elementary music education: A case study of S Elementary School in T City*. CNKI.
- Wang, Y. H. (2023). Exploring the effects of using various designs of game-based materials on music learning. *Interactive Learning Environments*, 31(5), 2650-2664.
- Yu, J. W. (2023). *Influence of tonal graphic notation on musical development of senior kindergarten children: A case study of D Kindergarten in S City*. CNKI.
- Zhou, D. Y. (2021). *Empirical research on Gordon's Music Aptitude Profile test in the Chinese context* [Master's thesis, Wenzhou University]. CNKI.