

Review of Evidence: Arts Education Through the Lens of ESSA

November 29, 2018

Yinmei Wan, Meredith J. Ludwig, and Andrea Boyle

American Institutes for Research



AMERICAN INSTITUTES FOR RESEARCH®

Review of Evidence: Arts Education Through the Lens of ESSA

November 29, 2018

Yinmei Wan, Meredith J. Ludwig, and Andrea Boyle



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Copyright © 2018 American Institutes for Research. All rights reserved.

This research was commissioned and funded by The Wallace Foundation as part of its mission to support and share effective ideas and practices.

<https://doi.org/10.59656/A-YA1991.001>

Acknowledgments

We express our gratitude to the staff of The Wallace Foundation, in particular Rachel Hare Bork, Ed Pauly, and Daniel Windham. Their questions, suggestions, and comments helped us improve the content and style of the report. We also wish to thank the many colleagues who helped us, especially Jim Lindsay, co-author of a preceding evidence review on arts integration, whose contributions to the design of the technical approach and analysis and synthesis of data guided the current review of arts education research. We appreciate the corrections and recommendations provided by Kathryn Young of EducationCounsel, who checked our presentation of the key elements in the Every Student Succeeds Act (ESSA) as they pertain to arts education. Three researchers at the American Institutes for Research (AIR) guided us in the content, technical approach and quality review for the report. David English informed our description of ESSA and the arts. Josh Polanin reviewed the meta-analysis work. Mengli Song provided overall quality review and feedback. Thank you to Elizabeth Scalia, our research librarian, who coached us in the development of the search parameters and made frequent suggestions, acquiring the materials that we needed for the screening and review.

Contents

	Page
Acknowledgments	iii
Executive Summary	1
Report Highlights	3
Recommendations.....	4
Concluding Thoughts	5
Chapter 1. What Do We Mean by “Arts Education?”	6
Focus of the Evidence Review	7
Arts in Education: Evolving Practice and Access	8
Research Literature About Arts in Education.....	11
Logic Model for Arts Education Interventions	12
Organization of the Remainder of This Report	13
Chapter 2. How Can the Every Student Succeeds Act Support Arts Education, and How Does It Define Evidence-Based Interventions?	15
Support for Arts Education Within ESSA.....	15
Identifying the Evidence Base of Arts Education Interventions	23
Chapter 3. What Evidence Exists Linking Arts Education Interventions With Improved Student Outcomes?.....	26
Evidence in Studies of Visual Arts Interventions	28
Evidence in Studies of Dance Interventions	32
Evidence in Studies of Drama Interventions.....	34
Evidence in Studies of Music Interventions	37
Evidence in Studies of Interventions That Involve Multiple Types of Art	46
Summary	50
Chapter 4. How Large Are the Effects of Arts Education Interventions on Student Outcomes?	52
Average Effect of Arts Education Interventions on Student Outcomes.....	55
Summary	60
References	61
Appendix A. Technical Approach to the Evidence Review	A-1
Appendix B. Evidence-Based Criteria for Grant Programs Awarded by the U.S. Department of Education	B-1

Tables

	Page
Table 2.1. Opportunities for Funding Under ESSA, Evidence Required for Funding, and Eligible ESSA Arts-Related Activities.....	17
Table 3.1. Evidence Review Results: Studies of Visual Arts Interventions With Tiers II and III Evidence.....	29
Table 3.2. Evidence Review Results: Studies of Visual Arts Interventions With Tier IV Evidence.....	30
Table 3.3. Evidence Review Results: Studies of Dance Interventions With Tiers II and III Evidence.....	33
Table 3.4. Evidence Review Results: Studies of Dance Interventions With Tier IV Evidence.....	33
Table 3.5. Evidence Review Results: Studies of Drama Interventions With Tier III Evidence.....	35
Table 3.6. Evidence Review Results: Studies of Drama Interventions With Tier IV Evidence.....	36
Table 3.7. Evidence Review Results: Studies of Music Interventions With Tiers II and III Evidence.....	39
Table 3.8. Evidence Review Results: Studies of Music Interventions With Tier IV Evidence.....	41
Table 3.9. Evidence Review Results: Studies of Interventions Involving Multiple Art Types With Tiers II and III Evidence.....	48
Table 3.10. Evidence Review Results: Studies of Interventions Involving Multiple Art Types With Tier IV Evidence.....	49
Table 4.1. Average Effects of Arts Education Interventions on Student Outcomes Within and Across Studies, by Art Type and Overall.....	56
Table 4.2. Average Effects of Arts Education Interventions, by Art Type and Type of Student Outcomes.....	57
Table A.1. Criteria for Determining Tier of Evidence Based on ESSA and U.S. Department of Education’s (2016) Non-Regulatory Guidance.....	A-7
Table A.2. Types of Information Recorded From Each Eligible Study Report.....	A-9

Figures

	Page
Figure 1.1. Components, Processes, and Outcomes: A Logic Model of Arts Education Interventions.....	12
Figure 2.1. “Evidence-Based” Intervention as Defined by ESSA.....	23

Figure 3.1. Evidence of the Effects of Arts Education Interventions on Student Outcomes: Number of Studies That Meet the Criteria for ESSA Evidence Tiers, by Art Type, Outcome Domain, and Tier of Evidence	27
Figure 4.1. Distribution of Improvement Indices for Intervention Effects on Student Achievement in Mathematics, Reading, and Science Based on WWC Intervention Reports	58
Figure A.1. Disposition of Reports of Studies of Arts Education Interventions Found in This Evidence Review	A-8
Figure A.2. Examples of Possible Effects of an Arts Education Intervention on Student Academic Achievement	A-11
Figure A.3. Computing and Interpreting the Improvement Index	A-12

Boxes

	Page
Box ES.1. Definitions of Key Terms	3
Box 1.1. Definitions of Key Terms.....	7
Box 1.2. Arts Education Practice: Evolving Definitions, Functions, and Key Players	10
Box 4.1. Meta-Analysis Explained	53
Box A.1. Extent of Literature Search.....	A-3

Executive Summary

Purpose of the Report

The Every Student Succeeds Act (ESSA), the 2015 reauthorization of the Elementary and Secondary Education Act, includes a diverse array of programs and funding streams that states, local educational agencies, and schools might leverage to support school improvement and student success. These programs and funding streams include approaches to teaching the arts and learning about the arts. Furthermore, ESSA contains provisions requiring or encouraging that educational agencies seeking to use federal funds available through the law for many of these approaches adopt evidence-based interventions. The inclusion of arts in ESSA shows that policymakers, not only arts advocacy groups and educators, view arts as an essential component in a well-rounded education.

This report presents the results of a review of evidence about arts education interventions based on the evidence requirements in ESSA. According to ESSA, for an intervention to be evidence based, research or theoretical support for the intervention must fall within one of four evidence tiers. Evidence in Tiers I–III must “demonstrate a statistically significant effect on improving student outcomes or other relevant outcomes,” and the three tiers represent varying levels of rigor (i.e., “strong/moderate/promising evidence”). Tier IV evidence must “demonstrate a rationale” that an intervention is “likely to improve student outcomes or other relevant outcomes,” and it must be coupled with “ongoing efforts to examine the effects” of the intervention (ESSA, Title VIII, Section 8101(21)(A)). School improvement activities funded through ESSA (Title I, section 1003) must include at least one intervention in one of the first three tiers.

Definition of Arts Education

The literature about arts education in prekindergarten through Grade 12 varies considerably in the use of terms to describe where and how students learn the knowledge and skills essential to an art type or discipline. For example, sometimes authors use the term “arts” to refer to all art types. When examining offerings in prekindergarten through Grade 12, the term “art” is often understood to refer to what we think of as “visual arts,” such as painting and drawing. We refer to this in our definition below. In the report we use the term “visual arts.”

For the purposes of this review, the term “arts education” is used and is defined as follows:

Arts lessons or classes offered in prekindergarten through Grade 12 that are (a) standards based and (b) taught by certified arts specialist teachers or teaching artists through (c) an explicit or implied sequential arts curriculum¹ in the (d) subjects of art/visual arts, media arts, music, dance, and drama/theater.²

¹ Sequential arts curriculum refers to courses or lessons designed to be coherent and taken in sequence in which students participate, starting at a beginning level and gradually moving toward a more advanced level of knowledge and skill.

² ESSA’s definition of a “well-rounded education” lists “the arts” and “music” as two distinct items. In this review, when the term “arts education” is used, music is understood to be one of the arts. Media arts can be considered a subject within arts; however, it is also often treated as its own art type.

Research Questions and Methods

The evidence review addresses two research questions:

1. Are there research studies on arts education interventions that meet the criteria for evidence as specified in ESSA?
2. How large are the effects of arts education interventions on student outcomes?³

To answer these two questions, we undertook five steps:

1. Informed by published research reports, arts organization websites, and other resources, we developed a logic model describing the features of arts education and developed search criteria that would be used in two research database searches.
2. We conducted two searches of educational research databases and clearinghouses to identify reports about the implementation and impact of arts education interventions. Both search efforts focused on research reported since 2000. The second search incorporated additional arts education terms to identify as many reports in scope as possible. The number of relevant reports found through these searches was 7,405.
3. Our search was refined as we screened abstracts and then full-text documents with a focus on finding empirical studies with data about student outcomes in prekindergarten through Grade 12. We screened out abstracts and texts deemed out of scope and identified 286 reports of studies for review by What Works Clearinghouse (WWC)–certified reviewers.
4. We examined the reports that contained studies of arts education interventions and classified the studies according to definitions of the ESSA tiers of evidence (Tiers I–IV).
5. Finally, to gain a fuller picture of the research findings on arts education interventions across all the well-designed studies, we recorded the magnitude of effects from 20 studies of arts education interventions and meta-analyzed those effect sizes.

Box ES.1 provides definitions to clarify the distinctions among some of the key terms used in this evidence review report.

³ Research question 2 originally expressed an interest in investigating the effects of arts education interventions particularly for students who are disadvantaged. The data reported in the studies included in this review, however, were not sufficient for this subgroup analysis.

Box ES.1. Definitions of Key Terms

Arts education intervention: A specific approach, set of activities, strategy, or program addressing the instruction, content, or outcomes of the instruction in and/or participation in the study of the arts, including art/visual arts, media arts, music, dance, or drama/theater.

Study: An empirical investigation of the effect of an arts education intervention on a particular sample and set of outcomes. Findings from a single study can appear in a single report or in multiple reports.

Report: A written summary of a study, in the form of a journal article, a book or book chapter, a dissertation, a technical report, or a conference paper. A report may present findings from a single study or multiple studies.

Report Highlights

ESSA addresses the issue of arts education in several ways. It maintains an emphasis throughout its varied funding streams on ensuring that students have access to a “well-rounded education,” which, according to the law, can include “the arts” and “music” along with other subjects. Title IV of ESSA explicitly identifies programs in the arts and arts integration as allowable activities, and it provides for dedicated Assistance for Arts Education program. ESSA also offers funding for arts teaching and learning interventions that address the needs of specific student subgroups, such as economically disadvantaged students and English learners.

ESSA includes more than 10 different funding opportunities that state educational agencies, local educational agencies, and schools can use to implement arts education interventions for students in prekindergarten through Grade 12. These funding opportunities can be used to support activities such as teacher professional development, school improvement efforts, supports for English learners, arts courses, instructional materials, and extended learning time programs. They can also be used to support arts-focused charter or magnet schools.

Evidence of the effects of arts education interventions on student outcomes exists at three ESSA evidence tiers. However, most arts education interventions included in this review are supported by Tier IV evidence only. This review identified a total of 88 studies with evidence in one or more of the four ESSA evidence tiers. Eighteen of the studies met the evidence requirements for Tiers II and/or III. The other 70 studies provided evidence in Tier IV only; that is, these studies demonstrate a theory- or research-based rationale for why the interventions studied should improve student outcomes (e.g., academic achievement, art learning, social-emotional learning, and process abilities such as creativity or critical thinking), but they show no statistically significant positive effect of the interventions on those outcomes.

The two art types with the largest number of studies with evidence in Tiers I-IV are music and visual arts. Fifty studies about music provided evidence aligned with ESSA tiers. The largest number of these studies, twenty-two, focused on art learning outcomes. Nineteen studies about visual arts provided evidence aligned with ESSA tiers. The largest number of these studies, ten, focused on process abilities as outcomes.

According to the meta-analysis conducted as part of this evidence review, the average effect found in the 20 well-designed studies examined was moderate and statistically significant (effect size = 0.38), indicating that an average child would gain 15 percentile points in a relevant student outcome examined in this review (i.e., academic achievement, art learning, social-

emotional learning, and process abilities) as a result of participating in an arts education intervention. The 15-percentile-point increase would put the average effect of arts education interventions at the 75th percentile among the 70 interventions in mathematics, reading, and science reviewed by the WWC. It is important to note the effects of arts education interventions included in this review varied by art type as well as by outcome type. Further, for some art types, there was as few as one intervention with statistically significant positive effects. Educators should be cautious in making decisions based on these meta-analysis results considering that such a small number of studies could influence the magnitude of impact.

Recommendations

Based on this evidence review, we offer several recommendations for practitioners and policymakers as well as researchers who are likely to engage in selecting, implementing, and evaluating interventions focused on arts education.

Recommendations for Practitioners and Policymakers

Be thoughtful in selecting which ESSA funding program(s) to pursue to support a proposed arts education intervention. Important factors to consider include the types of activities that are required and allowable under the program(s), the amount and duration of funding available, and the level of evidence required.

Critically assess the theoretical and empirical support behind a proposed arts education intervention. Adopting a conservative interpretation of ESSA's evidence-based criteria might help promote interventions with a stronger likelihood of success.

Before undertaking a search for arts education interventions that improve student outcomes, begin with a goal regarding the improvement focus as related to a particular art type and choose the intervention most likely to help reach that goal. As this evidence review notes, ESSA includes funding opportunities to use for arts education interventions and the literature contains many empirical reports of positive effects of such interventions on student outcomes. Still, practitioners and policymakers should be cautious when selecting interventions reporting positive evidence, as the studies providing such evidence may be limited in number, rigor, outcomes measured, and relevance to a particular context. Findings from reviews of studies by art type and by type of student outcomes may provide a more effective evidence base from which to consider implementing a particular arts education intervention.

Recommendations for Researchers

Research that is more rigorous is needed to provide stronger evidence for arts education. Researchers can help provide more Tier I evidence by using a randomized controlled trial study design, documenting the attrition of study participants, and providing sufficient details of analyses and findings in report appendices.⁴

⁴ In a randomized controlled trial, study participants are randomly assigned to an intervention group that receives the intervention or to a control group that does not receive the intervention.

Work toward a clear definition of the arts experience, a precise logic model, and appropriate instruments to measure desired outcomes of arts education interventions to achieve a high level of evidence. In *Critical Links: Learning in the Arts and Student Academic and Social Development* (Deasy, 2002), the reviewers of literature by art type consistently pointed out three challenges when examining the evidence about arts education interventions. The first is the existence of a description of the arts experience; the second is identifying the appropriate outcomes for the intervention; and the third is choosing the appropriate instruments to measure the outcomes. This evidence review has found similar omissions in studies conducted since 2000. Unfortunately, for most art types, except music, the literature continues to be sparse. The research literature on arts education interventions includes studies investigating the effects of arts education experiences on students' scores on standardized tests and college entrance exams. Other studies examine very specific art techniques such as teaching music composition, pitch, or self-evaluation while learning to play an instrument. Still others examine neurological changes in children, as a result of interventions that may or may not be adaptable to the school setting.

Additional research is needed to shed light on the effects of the use of arts education with diverse student populations in a range of settings. Researchers should plan their studies to include systematic comparisons of the effects of arts education on different student subgroups, such as students who are economically disadvantaged, English learners, and students with disabilities. Moreover, researchers should consider studying the effects of arts education among schools located in different settings (e.g., rural vs. suburban areas).

Concluding Thoughts

It may be timely to work toward a collaborative research agenda with emphasis on cross-cutting inquiry. In each decade since the 1990s, reviews of the condition of research about the arts have been undertaken and research agendas for the arts have been articulated. In 1994, the Office of Educational Research and Improvement and the National Endowment for the Arts sponsored a review of research and needs (National Endowment for the Arts, 1994). Deasy's (2002) *Critical Links* is another source of the condition of the literature and persisting research questions about arts education. Americans for the Arts and the American Educational Research Association sponsored the development of a research agenda (Arts Education Partnership, 2004). The Dana Foundation developed an agenda for research on the brain and the arts (Asbury & Rich, 2008). The work of examining the field and identifying knowledge gaps as well as a research agenda has been taken on by organizations representing professionals and educators in various art fields, such as the National Association of Music Merchants, the National Dance Education Association (Bonbright & Faber, 2004; Bradley, Bonbright, & Dooling, 2013), and the National Art Education Foundation. Over the past two years, the National Endowment for the Arts invested in research labs around priority topics, such as the arts and social/emotional well-being, innovation, and creativity and learning.

The identification of gaps from this evidence review may provide a jumping off point relevant for the future, as organizations consider a collaborative effort to examine the overarching questions that remain important to all the arts and could be pursued.

Chapter 1. What Do We Mean by “Arts Education?”

Chapter Highlights

- ▶ Arts education refers to arts lessons or classes offered in prekindergarten through Grade 12 that are (a) standards based and (b) taught by certified arts specialist teachers or teaching artists through (c) an explicit or implied sequential arts curriculum in the (d) subjects of art/visual arts, media arts, music, dance, and drama/theater.
- ▶ The practice of teaching about the arts in prekindergarten through Grade 12 has evolved with understanding about human development, technical and theoretical developments within each art discipline, and cultural and social influences, such as a shift in views about the importance of studying arts.
- ▶ The theory and research-based logic model for arts education interventions suggests that arts education interventions may focus on intermediary outcomes, such as teacher capacity, that may affect student outcomes. Arts education interventions may solely focus on student outcomes, such as art mastery, attitudes about art or arts education, social-emotional learning, and “process” abilities, such as creativity and critical thinking.

The 2015 reauthorization of the Elementary and Secondary Education Act as the Every Student Succeeds Act (ESSA) includes numerous federal funding opportunities that states, districts, and schools can use to support arts activities, strategies, and interventions designed to improve student learning. To help ensure the success of such interventions, the law requires or encourages that certain activities funded through ESSA be evidence based, and it outlines four evidence tiers for evaluating the level of evidence supporting a particular activity.

Evidence in Tiers I–III must “demonstrate a statistically significant effect on improving student outcomes or other relevant outcomes,” and the three tiers represent varying levels of rigor (i.e., “strong/moderate/promising evidence”). Tier IV evidence must “demonstrate a rationale” that an intervention is “likely to improve student outcomes or other relevant outcomes,” and it must be coupled with “ongoing efforts to examine the effects” of the intervention (ESSA, Title VIII, Section 8101(21)(A)).

This report presents findings from an evidence review of recent research on arts education interventions through the lens of ESSA’s four-tiered evidence framework. The review addresses two research questions:

1. Are there research studies on arts education interventions that meet the criteria for evidence as specified in ESSA?
2. How large are the effects of arts education interventions on student outcomes?⁵

⁵ This evidence review began with an additional goal to investigate the impact of arts education interventions particularly for students who are disadvantaged. The data reported in the studies included in this review, however, were not sufficient for this subgroup analysis. See Box 1.1 for definitions of key terms.

Box 1.1. Definitions of Key Terms

Arts education intervention: A specific approach, set of activities, strategy, or program addressing the instruction, content, or outcomes of the instruction in and/or participation in the study of the arts, including art/visual arts, media arts, music, dance, or drama/theater.

Study: An empirical investigation of the effect of an arts education intervention on a particular sample and set of outcomes. The findings from a single study can appear in a single report or in multiple reports.

Report: A written summary of a study, in the form of a journal article, a book or book chapter, a dissertation, a technical report, or a conference paper. A report may present findings from a single study or from multiple studies.

To address the above questions, the review team conducted a systematic search, screening, and review of reports of original studies published since 2000 that examined arts education interventions in prekindergarten through Grade 12 and their effects on student outcomes. In this chapter, we define the focus of the review, summarize the evolution of the practice of arts education, and describe characteristics of the arts education literature.

Focus of the Evidence Review

Prior to undertaking this evidence review, the review team conducted a pilot search of educational research databases and consulted a number of resources for clarity regarding the definition of arts education and the distinctions commonly made about art types or disciplines. Among the resources consulted were earlier, well-known research compendia (e.g., Deasy, 2002). The National Arts Standards (<https://www.nationalartsstandards.org/>) and websites for professional arts associations (e.g., National Dance Education Organization) were searched for sources of standards, art terms, and more recent statements about the status of research about arts education. In consultation with arts experts at The Wallace Foundation, we adopted the following definition of arts education for this evidence review:

Arts lessons or classes offered in prekindergarten through Grade 12 that are (a) standards based and (b) taught by certified arts specialist teachers or teaching artists through (c) an explicit or implied sequential arts curriculum in the (d) subjects of art/visual arts, media arts, music, dance, and drama/theater.⁶

This definition assumes the understanding that arts teaching and learning in prekindergarten through Grade 12 typically take place through instruction about concepts and skills related to a particular art form or discipline (e.g., art/visual arts, media arts, music, dance, or drama/theater). This is a distinguishing feature of arts education, as compared with arts integration, in that it takes place in a class or lesson devoted to the study of a particular art type. The definition also considers that the courses or lessons are designed to be coherent and taken in sequence, allowing participating students' knowledge and skills to develop towards a more advanced level.

⁶ ESSA's definition of a "well-rounded education" lists "the arts" and "music" as two distinct items. In our review, we consider music as one of the "arts" when we use the term arts education. While "visual arts" is sometimes specified in reports, occasionally the term "art" is used to refer to types of visual arts, such as painting, drawing, photography.

Arts in Education: Evolving Practice and Access

During the past 200 years, the content of arts education for each art type (i.e., what students study and produce) has expanded with input from professionals who practice that art and from teachers of the art (Szekely & Buckman, 2012). For example, the literature reviewers cited in *Critical Links* (Deasy, 2002) pointed to the evolution of the study of dance in public school, as having its origin in physical education classes (e.g., folk or square dance were often included in physical education). Similarly, reviewers of literature about drama education noted that prior to the existence of classes or clubs in theater, students were more likely to be engaged in the application of dramatic techniques in the study of literature (Deasy, 2002).

Arts are taught in various ways in the United States, which can vary based on the level of the school, the type of art, or the philosophy, and educational theories. The instructional setting for the study of arts could be a whole class, an individual, or small-group lessons (e.g., individual or group lessons for learning a particular type of instrument). In prekindergarten through Grade 12, teachers who lead classes in the arts are likely to be certified arts teachers. Arts therapy interventions, arts programs in summer camps, or afterschool arts programs were not within the scope of this evidence review. However, in the search for the literature, a number of studies of these out-of-scope interventions were located. For example, we identified studies of summer programs to prepare choral groups for upcoming competitions, studies of the use of arts to extend the educational opportunities for children or adults on the autism spectrum, and studies of a change in at-risk behavior in therapeutic or juvenile justice settings.

One new trend in arts education is the use of various technologies, such as software or online platforms, to teach the arts (Clements & Sarama, 2003; Peppler, 2013). Another trend is to study media arts as a course of study. Media arts may be structured as the study of a specific type of technique (e.g., computer-generated arts, or graphic arts) or a broad area of study that encompasses many techniques for a specific career-focused use (e.g., animation for film and television, advertising, journalism, museum education, and composing).

The question of access to arts education is one of the most common themes noted throughout the arts education literature. Evidence of an interest in showing the value of participating in arts education to students was noted recently when, at an Arts Education Partnership policy conference, state arts directors and arts educators highlighted efforts to identify a meaningful indicator about the arts in their ESSA-required state report cards. A recent report by the Brookings Institution called on arts educators to use available resources such as state and longitudinal databases to portray the extent of participation in the arts in their states (Kisida, Morrison, & Tuttle, 2017).

Differences in participation in arts classes or lessons were highlighted in the sampling and focus used for the 2016 National Assessment of Educational Progress arts assessment (NAEP Arts), which was administered only to students in the eighth grade, highlighting the limited participation across secondary grades. Moreover, NAEP Arts focused solely on music and visual arts because only a small percentage of schools offered students opportunities to study dance and theater. NAEP Arts found that the percentage of eighth graders in the United States taking music and visual arts classes or lessons stayed about the same in 2016 as in 2008: 63% of the students were enrolled in a class in music, and 42% of the students took a course in visual arts in both years

(National Center for Education Statistics, 2017). Perhaps in recognition of the statistics about access to arts classes, reduced funding, and the national NAEP Arts results, there is renewed understanding of the value of studying arts as part of an elementary, middle school, or high school curriculum. The inclusion of arts in ESSA shows that policymakers, not only arts advocacy groups and educators, view arts as an essential component in a well-rounded education.

The literature describing the history of arts in schools reflects advances in knowledge, political changes, and cultural changes, such as (a) an evolution in understanding how children learn; (b) increased value ascribed to education as a path to economic security; (c) changing expectations for student outcomes by the public and policymakers; (d) progressive, less formulaic approaches to curriculum and instruction; (e) shifting from an exclusive focus on the adult as artist to acceptance and value of the child as artist; (f) recognizing the importance of studying aesthetics and criticism; and (g) undertaking investigations into the relationship between different types of arts and other academic subjects (Fleming, 2012; National Endowment for the Arts, 1994; Szekely & Buckman, 2012). The resources from Fleming (2012), Szekely and Buckman (2012), and Stankiewicz, Amburgy, and Bolin (2004) are among many that point to some of the distinguishing features of various periods in the development of arts education in the United States. In Box 1.2, we summarize some of the notable events, people, and trends in arts education highlighted by these authors.

Box 1.2. Arts Education Practice: Evolving Definitions, Functions, and Key Players

Pre-Colonial America

- ▶ Families and societies passed down craft expertise.
- ▶ Explorers and missionaries brought traditions from their countries to the southwest.
- ▶ Artisans taught their craft to apprentices.

Colonial America (1700s)

- ▶ Tension existed between views of art's purpose: What is "high art" versus practical or mechanic arts?
- ▶ The study of the useful and/or the ornamental in art concerned primarily the education of White males since they were the only members of the population who owned property and were considered full citizens.
- ▶ Historians note an evolving view of the value of studying art techniques, history, aesthetics, and criticism for individual improvement.

Industrial America (1800s)

- ▶ Educators maintained that the American schoolrooms contributed the tools, knowledge, and skills relevant to future jobs.
- ▶ Art education was perceived as having special benefits for young women, and art and design schools were established as philanthropic enterprises in major cities.
- ▶ A major initiative to prepare teachers for drawing instruction occurred and was driven by educator Walter Smith.
- ▶ The artist professional community was developing.
- ▶ Drawing books and cards written by working artists were incorporated in schools, and the belief in drawing as a means of communicating and improving knowledge became a focus, leading to its establishment as a course in school.
- ▶ The Schoolroom Decoration and Picture-Study Movements focused on developing children's character by seeing and appreciating religious and moral meaning in adult art.
- ▶ Arthur Wesley Dow created the textbook *Composition*, explaining elements and principles of design.
- ▶ Manual training reflected interest in "practical" forms of education and reflected views about gender, race, and youth at risk.
- ▶ The National Education Association issued reports on American arts education.

Development of Theories of Instruction and Curriculum: Influences from other disciplines (1900s)

- ▶ The Child Study Movement emerged, which focused on understanding how children developed and learned and how to address problems.
- ▶ Psychologists viewed art as one way of studying the emotional health and cognitive growth of the child and as a therapeutic activity.
- ▶ Authors such as Viktor Lowenfeld and John Dewey produced theory-based works on arts education.
- ▶ After the 1950s, with a national preoccupation on math and science, art curriculum alternatives were designed to support their place in formal education.
- ▶ Getty Center for Education in the Arts seeded regional partnerships to implement balanced sequential arts instruction in the four areas of aesthetics, art criticism, art history, and art production.
- ▶ The theme of school reform through standards and accountability emerged and was maintained through A Nation At Risk, Goals 2000, and advocacy organizations.

Focus on Research and Dissemination (2000s)

- ▶ *Critical Links* (Deasy, 2002) and *Champions of Change* (Fiske, 1999) noted challenges for research on the arts.
- ▶ Arts in Education Model Development and Dissemination and Professional Development for Art Educators were established by the U.S. Department of Education.
- ▶ National Core Arts Standards are updated and disseminated.
- ▶ The National Center for Education Statistics (NCES) gathers data on the presence of arts education in public schools.
- ▶ NAEP Arts is a source of data about what eighth graders know and do in music and arts.
- ▶ Media technology becomes more present in schools at all levels, in software, as areas of study, and as the means to attract students to the arts.

Sources include: Stankiewicz, Amburgy, & Bolin, 2004, Fleming, 2012, and Szekely & Buckman, 2012.

Research Literature About Arts in Education

The drive to examine and improve the literature base about arts in education has intensified over the past 30 years. In each decade since the 1990s, reviews of the condition of research about the arts have been undertaken and research agendas for the arts have been articulated. The effect of this attention by collaborating national and professional organizations has moved the literature base from a primary focus on music and visual arts to an expanding literature base regarding dance and drama instruction and more rigorous research designs.

In 1994, the Office of Educational Research and Improvement and the National Endowment for the Arts sponsored a review of research and needs (National Endowment for the Arts, 1994). Deasy's (2002) *Critical Links* is another source of the condition of the literature and outstanding questions about arts education. Americans for the Arts and the American Educational Research Association sponsored the development of a research agenda on arts education (Arts Education Partnership, 2004). The Dana Foundation developed an agenda for research on the brain and the arts (Asbury & Rich, 2008). The Wallace Foundation has supported research about the quality of the arts experience (Seidel, Tishman, Winner, Hetland, & Palmer, 2009).

Other reports have come forward from work in the field by the National Association of Music Merchants (Hodges & Luehrsen, 2010), the National Dance Education Association (Bonbright & Faber, 2004; Bradley, Bonbright, & Dooling, 2013), and the National Art Education Foundation. Over the past two years, the National Endowment for the Arts invested in research labs around priority topics, such as social/emotional well-being, the arts and creativity, innovation, and learning.

The volume of published quantitative and qualitative studies of arts education has grown since 2000. Overall, there are indications that new and experienced researchers acknowledge the interest in more rigorous design for research on arts education and are bringing research techniques to bear with this goal in mind. Some of the more recent research has been conducted by Doctoral students, many of whom began as certified teachers of art and music, from U.S. colleges and universities through their dissertations. The research conducted for a number of dissertations follows the groundbreaking approach by Catterall (2009), that is, an analysis of existing longitudinal databases in which investigators examine the magnitude of the relationship between participation in arts classes and experiences and success at various stages of adult life.

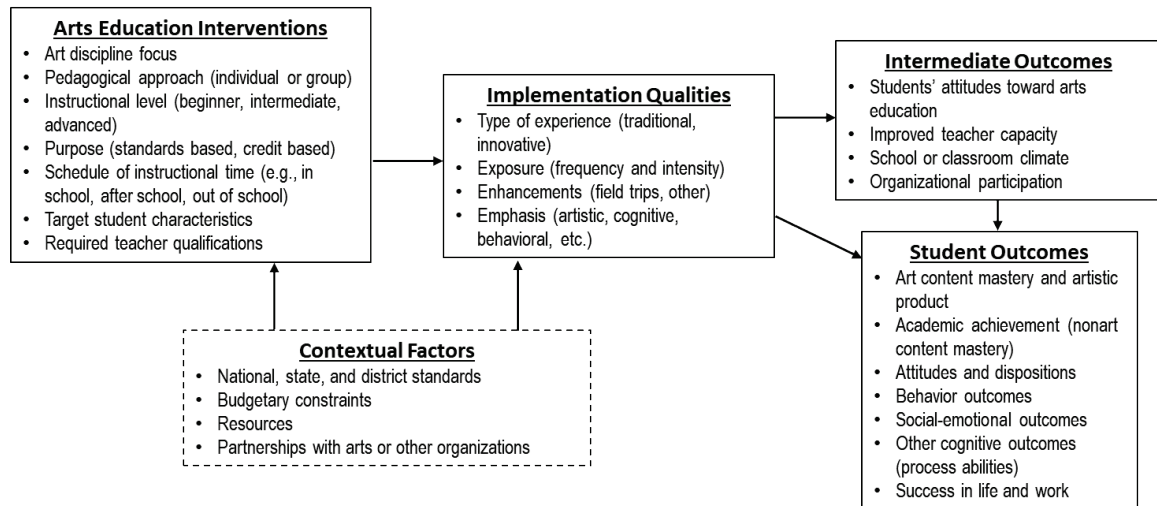
Among other arts educators and researchers, there has been a dialogue regarding the measurement of outcomes. Authors have noted the shifts in research focus, describing a determination to measure the art experience itself and to avoid proxy indicators such as academic achievement to measure the value of studying an art type (Fleming, 2012; Winner & Hetland, 2000). The literature contains studies of classroom conditions and technical skill development in a particular art form. Psychological attributes, such as motivation, creativity, and habits of mind, have also been investigated. The arena has expanded to consider the impact of participation in an arts-based learning experience for at-risk youth, for students with learning disabilities, and for students with economic disadvantages or second-language learning supports.

However, as highlighted in Deasy's (2002) *Critical Links*, questions persist as challenges for researchers today: Are experiences in different disciplines enough alike so they can be lumped together in arts education? Why can't we identify, define, and measure the collective multiple or integrated arts learning experience very well? What are the limitations of studies that measure arts learning using engagement or class participation as an indicator of improving outcomes? What methods are available to measure learning outcomes directly, whether these are creativity, self-concept, or performance?

Logic Model for Arts Education Interventions

In the literature, arts education interventions are theorized to impact student outcomes. The logic model as depicted in Figure 1.1 captures some of the key features, activities, and outcomes noted in the literature.

Figure 1.1. Components, Processes, and Outcomes: A Logic Model of Arts Education Interventions



The box on the far left of the logic model represents **arts education interventions** as they were studied, and it describes some of the key features of these interventions. Arts education interventions, for example, typically include an emphasis on an art type or discipline (e.g., visual arts). Additional intervention features may include specific pedagogical approaches, instructional level, and credit-bearing status. Targeted student participation characteristics and required teacher qualifications may vary as well as scheduling. Arts education interventions may address the learning needs of different student populations, such as students with special learning needs or students of a particular grade level.

The **implementation qualities** box identifies dimensions along which arts education interventions may vary. Arts education interventions can differ in instructional approach and content coverage: Some programs describe themselves as traditional, as in teacher-led lessons, and some interventions describe themselves as innovative, where students take a leadership role. The exposure of students to different interventions may vary in frequency and intensity. Some of the interventions include enhancements to regular classes such as field trips. Finally, approaches may

differ in their emphasis, that is, whether the focus is on creating art products, learning to think critically, or using studies in drama to address themes relevant to school behaviors, such as bullying.

The implementation of arts education interventions may lead to improved **intermediate outcomes** that are expected ultimately to contribute to other key student outcomes. Examples of intermediate outcomes at the student level include students' attitudes toward an art type or toward arts education in general. Some interventions focus on teacher practice, such as teachers' ability to employ specialized instructional strategies, classroom interactions and behavior/motivation techniques, and assessment practices. Intermediate outcomes may include school-, parent-, or community-level measures such as school climate as well as the participation of partner organizations, such as museums.

These intermediary outcomes may, in turn, affect key **student outcomes** such as art content mastery (e.g., art achievement) and the creation of art products (e.g., performance); academic achievement (i.e., nonart content mastery); attitudes and dispositions; behavior outcomes (e.g., attendance); social-emotional skills (e.g., self-confidence and self-awareness); other cognitive outcomes (i.e., process abilities such as creativity and critical thinking); and success in life and work.

Contextual factors related to the circumstances in which an arts education intervention takes place can influence the intervention's design as well as its implementation and outcomes. For example, a school district's attention to recommendations from professional arts organizations or national standards regarding sequencing art classes may affect offerings in schools and expectations for student outcomes. Budgetary constraints may affect exploration of approaches and innovations such as the use of new technologies. The knowledge, skills, and dispositions of participating teachers might determine how much and what types of teacher support are included in an intervention's design. Moreover, the availability of other resources that are important for sustaining an intervention over time (e.g., community support and school leadership buy-in) can play key roles in an intervention's design and implementation. Similarly, features of the instructional setting (e.g., type of physical space available and proximity to museums or performing arts centers) and the availability of instructional resources (e.g., necessary technology and art supplies) can shape how an arts education intervention is devised and carried out. Finally, it is possible that through the partnership with local or national arts support organizations, the arts education program in a school or throughout a district can be started and sustained.

Organization of the Remainder of This Report

In Chapter 2, we describe the funding opportunities available in ESSA for arts education interventions as well as the specific evidence requirements and guidelines for determining whether those interventions meet ESSA's definition of evidence based. The findings from our evidence review are presented in Chapters 3 and 4. Chapter 3 identifies studies of arts education interventions that met the criteria for ESSA Evidence Tiers I–IV by art type. Chapter 4 examines the magnitude of the effects of arts education interventions on student outcomes. A list of references follows Chapter 4. Appendix A provides detailed information about the review methods. Appendix B describes the criteria established by the U.S. Department of Education for each evidence tier for use in grant competitions that are administered directly by the Department. Finally, a large number of studies in this review of arts education interventions did

not provide evidence aligned with the requirements for Tiers I–IV. Rather than inserting multiple pages with this list in this report, we have created a separate listing that is available at the same location as the report.

Chapter 2. How Can the Every Student Succeeds Act Support Arts Education, and How Does It Define Evidence-Based Interventions?

Chapter Highlights

- ▶ ESSA offers over 10 different funding opportunities that state educational agencies, local educational agencies, and/or schools can use to support arts education.
- ▶ Although ESSA's definition of evidence-based interventions provides some criteria for evaluating the level of evidence supporting a particular intervention, it leaves room for interpretation. Interpretation for this evidence review was informed by prior evidence reviews and the expertise of colleagues at AIR and EducationCounsel.
- ▶ This evidence review adopts the evidence criteria specified in the Non-Regulatory Guidance issued by the U.S. Department of Education in September 2016, which are stricter than the criteria specified in ESSA.

In this chapter, we consider the decisions that practitioners and policymakers will make as they plan to use arts education interventions supported by funding under ESSA. First, we explore which funding programs within ESSA may lend themselves to supporting particular types of arts education interventions based on the programs' stated purposes and allowable activities. We then unpack ESSA's definition of "evidence based" interventions and explain how we applied that definition to our evidence review.

Support for Arts Education Within ESSA

ESSA addresses the issue of arts education in several important ways. For one, it broadens the subject area emphasis on mathematics, reading/English language arts, and science compared with its predecessor, the No Child Left Behind Act, to include a "well-rounded education." Although the law does not dictate what subjects constitute a well-rounded education, it clarifies that a well-rounded education can include "the arts" and "music" along with other academic subject areas, such as history and foreign languages (ESSA Section 8101(52)). An increased focus on ensuring that students have equitable access to a well-rounded education runs throughout ESSA's various titles and thus opens the door to using numerous ESSA funding programs to help finance arts education interventions.

In addition to supporting the arts as part of a well-rounded education, ESSA features a funding stream devoted specifically to advancing arts education. Title IV, Part F of ESSA establishes a dedicated Assistance for Arts Education program (Sec. 4642 of ESSA, similar to the No Child Left Behind Act's Arts in Education program), which supports a range of activities to promote education in music, dance, theater, media arts, and visual arts. State and local educational agencies can use Assistance for Arts Education funding for activities such as providing professional development to arts educators, developing instructional programs and materials in the arts, and expanding partnerships with museums, centers for the arts, and other arts organizations.

Last, ESSA provides funding opportunities to increase support for particular student subgroups such as students who are economically disadvantaged and English learners. Those opportunities can be used to fund interventions that help arts educators address those subgroups' specialized learning needs. ESSA also requires states to set aside funds for school improvement activities in their lowest-performing schools, which could support the use of school turnaround strategies that incorporate sequential arts instruction, such as an arts-focused school improvement model or a school improvement model that includes arts instruction as part of a well-rounded education.

To qualify for funding under a specific ESSA grant program, an arts education intervention must align with the grant program's purpose and allowable activities. It also must satisfy any applicable evidence-based intervention requirements associated with that program. For certain ESSA programs, such as Title I, Section 1003: School Improvement, the law explicitly requires the use of evidence-based interventions for at least some of the activities allowable under that program. For other programs, the law indicates that evidence-based interventions must receive competitive preference in grant competitions, meaning that grant applications are not required to propose evidence-based interventions but must have a greater chance of being awarded if they do. Finally, some ESSA grant programs do not specify any evidence-based intervention requirements. Nevertheless, applicants may still choose to propose evidence-based interventions for these programs in the hopes, for example, that such interventions will be more likely to improve student outcomes.

Table 2.1 presents a summary of ESSA funding opportunities that seem well suited for supporting arts education interventions, together with a summary of the applicable evidence-based intervention requirements for each opportunity listed.

Table 2.1. Opportunities for Funding Under ESSA, Evidence Required for Funding, and Eligible ESSA Arts-Related Activities

Opportunities for Supporting Arts Education	Description	Eligible Entities	Authorized Funding ^a	Evidence Required ^b	Examples of Eligible Arts Education Activities ^c
Title I, Section 1003: School Improvement ^d	Funds are intended to support improvement activities and technical assistance for schools identified for comprehensive or targeted support and improvement under their state accountability system. ^e	LEAs with low-performing schools identified for comprehensive or targeted support and improvement by their state’s accountability system	SEAs must reserve 7% of their Title I, Part A allocation.	At least one intervention must meet evidence requirements of Tier I, II, or III.	<ul style="list-style-type: none"> ▶ Arts-focused school turnaround models that feature arts instruction ▶ Professional development for arts educators in schools identified for comprehensive or targeted support and improvement ▶ Incorporating arts into a school improvement plan as a means of promoting a well-rounded education
Title I, Section 1003A: Direct Student Services	<ul style="list-style-type: none"> ▶ Funds may be used to support academic coursework not otherwise available at the school, as well as personalized learning, all aimed at improving college and career readiness and success. ▶ Examples include advanced courses, career and technical education courses, credit recovery or academic acceleration courses, and personalized learning approaches. 	LEAs that reflect geographic diversity within the state, with priority given to LEAs serving the highest percentage of schools identified for comprehensive or targeted support and improvement	SEAs may reserve up to 3% of their Title I, Part A allocation.	No minimum evidence threshold	<ul style="list-style-type: none"> ▶ Arts courses that support schools’ provision of a well-rounded education ▶ Advanced Placement, International Baccalaureate, and postsecondary level courses in the arts (and for low-income students, costs associated with exams that provide postsecondary credit in the arts) ▶ Courses that lead to industry-recognized credentials in the arts

Opportunities for Supporting Arts Education	Description	Eligible Entities	Authorized Funding ^a	Evidence Required ^b	Examples of Eligible Arts Education Activities ^c
Title I, Part A: Basic Programs for Schoolwide and Targeted Assistance ^d	<ul style="list-style-type: none"> ▶ Funds are intended to support schools and LEAs that serve high numbers or percentages of students from low-income families. ▶ Schools operating Title I schoolwide programs can use their funds to improve the school's entire educational program. ▶ Schools operating targeted assistance programs must focus Title I-funded activities on students who are failing or most at risk of failing to meet state standards. 	Schools where at least 40% of students come from low-income families (and other schools that receive a waiver from the state) qualify for Title I schoolwide programs; schools with smaller low-income populations can implement targeted assistance programs.	\$15 billion for FY 2017 to \$16.2 billion for FY 2020	No minimum evidence threshold, except that any external providers selected to help schools implement their schoolwide or targeted assistance programs must have expertise in using evidence-based strategies (Tiers I–IV).	<ul style="list-style-type: none"> ▶ Whole-school improvement models or strategies that feature arts instruction ▶ Professional development to support arts educators ▶ Arts education programs or courses that support students' access to a well-rounded education ▶ Arts programs that promote parent engagement
Title II, Part A: Supporting Effective Instruction ^d	Funds are intended to recruit and enhance the quality and effectiveness of current teachers, principals, and other school leaders.	LEAs	Approximately \$2.3 billion per year for FY 2017 to 2020	Some specific uses of funds (e.g., professional development, induction, and mentoring) require Tier I, II, III, or IV evidence, to the extent that the state determines such evidence is reasonably available. ^f	<ul style="list-style-type: none"> ▶ Teacher professional development activities for arts educators ▶ Collaboration time for teaching artists and arts educators to plan lessons ▶ Time for teachers to develop curricula in the arts ▶ Financial incentives to recruit arts educators ▶ Financial incentives to help qualified individuals with an arts background become art educators

Opportunities for Supporting Arts Education	Description	Eligible Entities	Authorized Funding ^a	Evidence Required ^b	Examples of Eligible Arts Education Activities ^c
Title III, Part A: English Language Acquisition, Language Enhancement, and Academic Achievement Act ^d	Supplemental funding provides support for ELs in attaining English proficiency and developing high levels of academic achievement.	LEAs or consortia of multiple LEAs that serve sufficient numbers of ELs	Approximately \$760 million for FY 2017 to \$885 million for FY 2020	No minimum evidence threshold	<ul style="list-style-type: none"> ▶ Professional development to build arts educators' capacity to meet the needs of ELs ▶ Acquiring digital resources that support ELs in arts classes
Title IV, Part A: Student Support and Academic Enrichment Grants ^d	Funds provide support for: <ul style="list-style-type: none"> ▶ providing all students with access to a well-rounded education which may include the arts, ▶ improving school conditions for student learning to create a healthy and safe school environment, and ▶ effectively using technology to improve academic achievement and digital literacy. 	LEAs or consortia of multiple LEAs	Approximately \$1.6 billion per year	Some specific uses of funds require Tier I, II, III, or IV evidence, to the extent that the state determines such evidence is reasonably available. ^d	<ul style="list-style-type: none"> ▶ Arts instructional programs to provide students with a well-rounded education ▶ Arts programs that promote students' physical and mental health, trauma recovery, or social-emotional skills (e.g., resilience, perseverance, self-management, empathy) ▶ Professional development and equipment to support the effective use of technology in delivering arts instruction, supporting collaboration among arts educators, and/or engaging with arts professionals outside of the school
Title IV, Part B: 21st Century Community Learning Centers	Funds are intended to support academic enrichment activities and other support services that supplement regular school day requirements.	LEAs or other public or private nonprofit entities, with priority given to those who serve students from high-poverty and low-performing schools	\$1 billion for FY 2017 and \$1.1 billion per year from FY 2018 to 2020	Tier I, II, III, or IV evidence, when deemed appropriate	<ul style="list-style-type: none"> ▶ Expanded learning time (e.g., before- or after-school programs) that provide arts instruction ▶ Summer school programs that provide arts instruction

Opportunities for Supporting Arts Education	Description	Eligible Entities	Authorized Funding ^a	Evidence Required ^b	Examples of Eligible Arts Education Activities ^c
Title IV, Part C: Expanding Opportunity Through Quality Charter Schools	Start-up funds are provided for creating new charter schools or replicating charter schools with a track record of success.	Charter management organizations, with priority given to those that serve high-poverty student populations	\$270 to \$300 million per year for FY 2017 to 2020	No minimum evidence threshold	<ul style="list-style-type: none"> ▶ Planning activities for establishing or replicating charter schools with a specialized arts focus or curriculum ▶ Professional development for teachers working in new charter schools to help them provide arts instruction
Title IV, Part D: Magnet School Assistance	Start-up funds are provided for establishing public schools or education centers that offer a specialized, theme-based instructional program designed to attract and bring together students from different racial and socioeconomic backgrounds.	LEAs or consortia of multiple LEAs that serve students from diverse racial and socioeconomic backgrounds	\$94 million for FY 2017 to \$108 million for FY 2020	Competitive preference is given for proposals with evidence-based activities (Tier I, II, III, or IV).	<ul style="list-style-type: none"> ▶ Creating or replicating a magnet school with a specialized arts focus or curriculum ▶ Professional development for teachers working in a new magnet school to help them provide arts instruction
Title IV, Part F: National Activities Subpart 1— Education Innovation and Research	Funds are intended to support the development, implementation, or scale-up of innovations to improve high-need students' achievement and attainment, along with rigorous evaluations of those innovations.	SEAs; LEAs; Bureau of Indian Education; nonprofits; consortia of SEAs or LEAs; or partnerships between SEAs, LEAs, or Bureau of Education and nonprofits, businesses, educational service agencies, or institutions of higher education	\$70.5 million for FY 2017 to \$90.6 million for FY 2020	Includes program-specific evidence requirements	<ul style="list-style-type: none"> ▶ Developing and evaluating a field-initiated arts teaching or learning intervention ▶ Scaling up and rigorously evaluating an arts teaching or learning intervention with positive prior evaluation results to increase its evidence base

Opportunities for Supporting Arts Education	Description	Eligible Entities	Authorized Funding ^a	Evidence Required ^b	Examples of Eligible Arts Education Activities ^c
<p>Title IV, Part F: National Activities Subpart 2—Community Support for School Success</p>	<ul style="list-style-type: none"> ▶ Promise Neighborhoods grants support the provision of comprehensive, coordinated services to neighborhoods with high rates of poverty, multiple signs of distress (e.g., high rates of academic failure, obesity, incarceration), and low-performing schools. ▶ Full-Service Community Schools grants support the coordination and provision of pipeline services (a continuum of coordinated services from birth to postsecondary education and career attainment) in public elementary or secondary schools. 	<p><i>Promise Neighborhoods:</i> Non-profits, institutions of higher education, Indian tribes</p> <p><i>Full-Service Community Schools:</i> LEAs working in partnership with one or more community-based organizations, nonprofit organizations, or other public or private entities</p>	<p>Approximately \$70.5 million per year for FYs 2017 and 2018 and approximately \$70 million per year for FYs 2019 and 2020</p>	<p><i>Promise Neighborhoods:</i> Application requirement includes support for evidence-based programs that assist students through transitions (e.g., pre-K–12 to postsecondary), and competitive preference is given for proposals with evidence-based activities (Tier I, II, III, or IV).</p> <p><i>Full-Service Community Schools:</i> Competitive preference is given for proposals with evidence-based activities (Tier I, II, III, or IV).</p>	<ul style="list-style-type: none"> ▶ Arts initiatives that focus on improving students’ academic and developmental outcomes (e.g., school readiness, social-emotional skills, mental and physical health) ▶ Arts programs for early childhood students ▶ Extended learning time opportunities (e.g., before- or after-school or summer programs) that provide arts instruction ▶ Partnerships with community arts organizations to support arts initiatives

Opportunities for Supporting Arts Education	Description	Eligible Entities	Authorized Funding ^a	Evidence Required ^b	Examples of Eligible Arts Education Activities ^c
Title IV, Part F: National Activities Subpart 4—Awards for Academic Enrichment	Funds include Assistance for Arts Education Grants that support efforts to promote arts education, including arts education for students who are disadvantaged and students with disabilities.	<ul style="list-style-type: none"> ▶ An LEA or a consortium of LEAs where at least 20% of students come from low-income families ▶ SEAs ▶ Institutions of higher education ▶ Museums or cultural institutions ▶ Bureau of Indian Education ▶ Other nonprofit or private organizations 	Approximately \$55 to \$56 million per year for FY 2017 to 2020	No minimum evidence threshold	<ul style="list-style-type: none"> ▶ Arts courses or programs ▶ Professional learning programs for art educators ▶ Development and use of instructional materials and digital resources to support arts instruction ▶ Partnerships with arts organizations, museums, or performing arts centers

Note. EL = English learner. FY = fiscal year. LEA = local educational agency. SEA = state educational agency.

^a The funding levels reported reflect the total amount of funding authorized in the law; they do not necessarily reflect the actual funding levels for each year, which are determined through the annual federal budget process.

^b This column reflects evidence requirements that are outlined in an ESSA statute. For grant competitions that are conducted directly by the U.S. Department of Education (rather than by SEAs), the Department might choose to add a competitive preference for evidence-based interventions even if not explicitly required under the law.

^c The activities listed are only examples of eligible arts activities for these programs as provided in statute, nonregulatory federal guidance, or otherwise; it is not an exhaustive list.

^d Indicates that states distribute funds to eligible entities based on a formula rather than on a competitive basis. For Title I 1003A Direct Student Services funds and Section 1003 School Improvement funds, states have the option of issuing subgrants to LEAs on either a formula or competitive basis.

^e Comprehensive support and improvement schools are schools among the lowest 5% of Title I schools across all required indicators within their state's accountability system, schools that fail to graduate one third or more of their students, and Title I schools with chronically underperforming student subgroups. Targeted support and improvement schools are schools with one or more student subgroups that are consistently underperforming, as defined by the state, based on all required indicators within their state's accountability system. Schools with one or more subgroups that would perform, on their own, as poorly as the lowest 5% of Title I schools receive additional targeted support and improvement activities.

^f In some instances, ESSA allows states to waive evidence-based intervention requirements for specific uses of funds if the state determines that an evidence base is not reasonably available to apply to those requirements.

Identifying the Evidence Base of Arts Education Interventions

Title VIII, Section 8101 of ESSA defines four tiers of evidence for evaluating the level of rigor in the research base for evidence-based interventions, as shown in Figure 2.1. Interventions supported by evidence in Tiers I–III must “demonstrate a statistically significant effect on improving student outcomes or other relevant outcomes,” and the three tiers represent varying levels of rigor (i.e., “strong, moderate, or promising evidence”). An intervention supported by Tier IV evidence must “demonstrate a rationale” that it is “likely to improve student outcomes or other relevant outcomes,” and it must be coupled with “ongoing efforts to examine the effects” of the intervention.

Figure 2.1. “Evidence-Based” Intervention as Defined by ESSA

WHAT IS AN “EVIDENCE-BASED” INTERVENTION?
(from section 8101(21)(A) of the ESEA)

“...the term ‘evidence-based,’ when used with respect to a State, local educational agency, or school activity, means an activity, strategy, or intervention that –

- (i) demonstrates a statistically significant effect on improving student outcomes or other relevant outcomes based on –
 - (I) *strong evidence* from at least one well-designed and well-implemented experimental study;
 - (II) *moderate evidence* from at least one well-designed and well-implemented *quasi-experimental study*; or
 - (III) *promising evidence* from at least one well-designed and well-implemented correlational study with statistical controls for selection bias; or
- (ii) (I) *demonstrates a rationale* based on high-quality research findings or positive evaluation that such activity, strategy, or intervention is likely to improve student outcomes or other *relevant outcomes*; and
 - (II) includes ongoing efforts to examine the effects of such activity, strategy, or intervention.

As Table 2.1 in the previous section shows, under most federal programs that require or encourage the use of evidence-based interventions, interventions can satisfy ESSA’s definition of evidence-based if they meet the requirements for any of the four evidence tiers. However, the law requires that school improvement activities funded under Title I, Section 1003, must include at least one intervention that meets the evidence requirements for one of the first three (i.e., the most rigorous) tiers.

Although ESSA’s definition of evidence-based interventions outlines general criteria for each evidence tier, it leaves room for states to specify more detailed criteria in areas in which the law is silent (Herman et al., 2016).⁷ In this section, we outline the general approach that we took in interpreting and applying ESSA’s evidence-based intervention criteria for the purposes of this evidence review, which was informed by prior evidence reviews through the lens of ESSA and the expertise of colleagues knowledgeable about ESSA, the regulations, the guidance, and specific program requirements within ESSA. However, we recognize that guidance and interpretations of ESSA’s new requirements for

⁷ The U.S. Department of Education has established its own criteria for each evidence tier for use in grant competitions that are administered directly by the Department (see Appendix B).

evidence-based interventions are still evolving as of this writing, and thus we encourage the reader to delve more deeply into available guidance and other resources on the evidence-based intervention criteria to make informed decisions.

For this evidence review on arts education interventions, we adopted the stricter evidence criteria laid out in a September 2016 U.S. Department of Education document that provides nonregulatory guidance⁸ to states and school districts on how they might interpret the four evidence tiers (U.S. Department of Education, 2016). This document suggests that evidence tiers be defined using the following general criteria:

- ▶ Strong evidence (Tier I) comes from study reports that (a) show statistically significant positive intervention effects on relevant outcomes (without any statistically significant negative effects); (b) meet What Works Clearinghouse (WWC) evidence standards without reservations⁹; and (c) were conducted using a large, multisite sample (i.e., more than 350 students and more than a single school district).
- ▶ Moderate evidence (Tier II) comes from study reports that (a) show statistically significant positive intervention effects on relevant outcomes (without any statistically significant negative effects); (b) describe studies that meet WWC evidence standards with reservations; and (c) describe studies that were conducted using a large, multisite sample (i.e., more than 350 students and more than a single school district).
- ▶ Promising evidence (Tier III) comes from study reports that (a) show statistically significant positive intervention effects on relevant outcomes (without any statistically significant negative effects), and (b) describe correlational studies with statistical controls for selection bias.
- ▶ Research-based rationale (Tier IV) evidence comes from study reports that (a) feature a well-specified logic model informed by research or evaluation, and (b) describe interventions that are undergoing additional research regarding their effects.

As Herman et al. (2016) noted, the criterion for Tier IV evidence—a “research-based rationale”—is particularly challenging to apply because ambiguities in the definition of Tier IV evidence leave it open to broad or narrow interpretation. In their review of school leadership research, Herman et al. (2016) drew on information from the U.S. Department of Education’s September 2016 Non-Regulatory Guidance to interpret the first criterion for Tier IV evidence—a “well-specified logic model informed by research or evaluation”—to mean (a) a graphically presented logic model that includes key components of the intervention and outcomes, where (b) research or evaluation findings exist to support a connection between at least one intervention component and at least one relevant outcome (U.S. Department of Education, 2016, p. 9).

Because the study reports collected for this review of arts education interventions vary considerably in how they present rationales or theories for the ways in which the interventions are expected to

⁸ Non-Regulatory Guidance is information released by the U.S. Department of Education to assist SEAs and LEAs in implementing particular provisions of the law. It might include explanations, examples, and suggestions on how to implement the law’s provisions, but it does not require the use of specific activities or practices.

⁹ Details about study designs that meet WWC evidence standards with or without reservations can be found in WWC’s standards handbook. This evidence review was guided by WWC’s Version 3.0 standards, which, together with the latest Version 4.0 standards, can be found at: <https://ies.ed.gov/ncee/wwc/Handbooks>.

influence student outcomes, we decided to employ a more inclusive approach to applying Tier IV evidence requirements. We considered studies that featured any type of rationale or theory of action linking features of the intervention to relevant outcomes as candidates for Tier IV evidence. We did not evaluate whether the studies met the second criterion for Tier IV evidence (i.e., that they describe an intervention that is undergoing additional research) as part of this review.¹⁰ As we examined the studies that were candidates for Tier IV evidence, we noted that although an explicit logic model was absent in many of the studies, most included research references which the authors maintained were key to the intervention components and hypotheses guiding their investigations. This made sense to us, since the studies were published in peer-reviewed journals and/or were theses or dissertations. Conventionally, these kinds of publications often include a review of the literature related to the intervention investigated. In this report, unless there is an existing logic model in a study, we refer to the Tier IV evidence as being either a theory or research-based rationale.

For this review, we classified studies as providing *no tier-aligned evidence* if the following two conditions were met: (a) The study lacked statistically significant findings or used a research design other than those eligible for providing evidence in Tiers I–III and (b) the report describing the study lacked any type of rationale or logic model for the intervention studied. More details about the methodology used to identify studies on arts education interventions and classify the evidence from those studies according to the ESSA evidence tiers are provided in Appendix A.

¹⁰ Due to our more inclusive approach regarding studies of interventions with Tier IV evidence, the interventions we identified as providing Tier IV evidence may require additional scrutiny to determine whether they satisfy the Department's criteria for Tier IV evidence.

Chapter 3. What Evidence Exists Linking Arts Education Interventions With Improved Student Outcomes?

Chapter Highlights

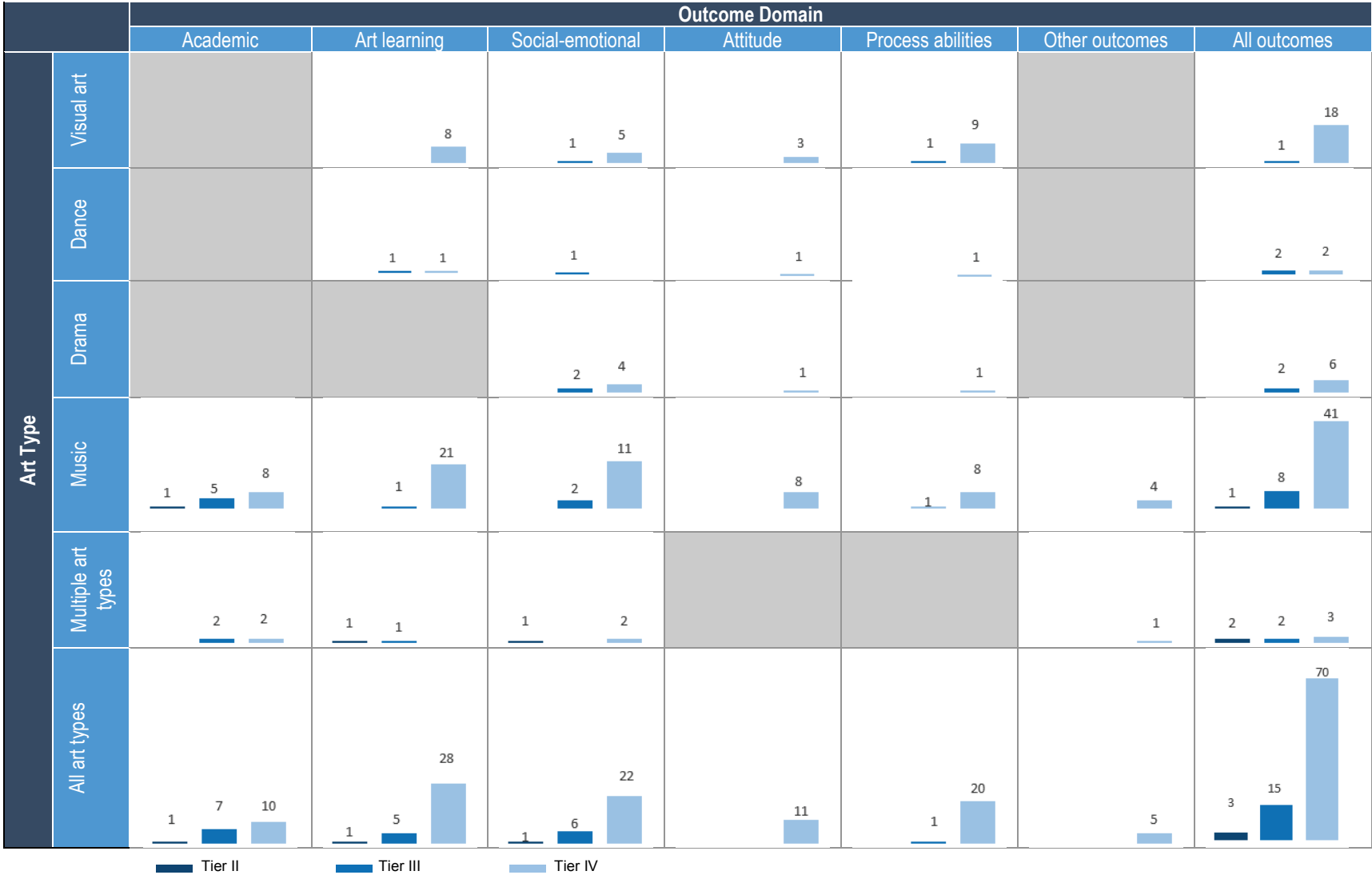
- ▶ Evidence of the effects of arts education interventions on student outcomes exists at three of ESSA's evidence tiers.
 - Three studies provided evidence at Tier II and 15 studies provided Tier III evidence regarding the type of arts education interventions within the scope of this evidence review.
 - Seventy studies provided evidence at Tier IV for arts education interventions.
 - No studies of arts education interventions provided Tier I evidence.

The findings from the evidence review are divided into two chapters. Chapter 3 presents findings for the first research question (repeated below), and Chapter 4 presents the meta-analytic findings for the second research question.

Research question 1: Are there research studies on arts education interventions that meet the criteria for evidence as specified in ESSA?

To address this question, we classified the evidence from each relevant study on arts education interventions based on whether the study and its findings met the criteria for ESSA's Evidence Tiers I–IV. Figure 3.1 summarizes the number of studies providing evidence at one of the four tiers, overall and by type of art and type of outcome. The bars in each cell indicate the numbers of studies providing evidence at Tiers II, III, and IV respectively. Empty cells (shaded grey) indicate that no study reviewed provides evidence aligned with ESSA tiers for that particular art type and outcome domain. The details about each study, its findings, and the reason for its classification are provided in this chapter in sections devoted to the art type.

Figure 3.1. Evidence of the Effects of Arts Education Interventions on Student Outcomes: Number of Studies That Meet the Criteria for ESSA Evidence Tiers, by Art Type, Outcome Domain, and Tier of Evidence



Note. Shaded cells indicate that no study reviewed provides evidence aligned with ESSA tiers for that particular art type and outcome domain.

Source: Authors' analysis of findings from literature review on arts education interventions.

As Figure 3.1 shows, evidence of the effects of arts education interventions on student outcomes exists at ESSA evidence Tiers II, III, and IV. No interventions were supported by ESSA Tier I evidence. A total of 88 studies provided evidence meeting ESSA evidence criteria for 87 arts education interventions¹¹: three studies with Tier II evidence, 15 studies with Tier III evidence, and 70 studies with Tier IV evidence. The number of studies of interventions in the arts of music (50 studies) and visual arts (19 studies) far exceeds that for the arts of drama (eight studies) and dance (four studies).

Most studies of arts education interventions included in this review provided Tier IV evidence only. Reports providing this type of evidence lack empirical results from rigorous studies, but they do contain a theory- or research-based rationale for why the intervention should improve student outcomes. Two studies provided evidence that could be classified at both Tier II and Tier III; 10 studies provided evidence that could be classified at both Tier III and Tier IV. These studies are represented just once—in the count for the higher level of evidence—within each cell of Figure 3.1. Nineteen studies reported evidence in multiple outcome domains that could be classified at Tiers II, III, or IV.

In the remainder of this chapter, we present evidence supporting arts education interventions by type of art. There are two tables for each art type: One table lists studies of interventions with evidence at Tiers II and III, and the second table lists studies of interventions with evidence at Tier IV. Because the Tier IV table includes many more studies, it lists the studies of interventions alphabetically by leading author within each grade span for ease of locating the studies within the reference list.

Evidence in Studies of Visual Arts Interventions

Our review found 19 studies of visual arts interventions that provided evidence at Tiers III and IV. One study produced evidence aligned with Tier III (Table 3.1), and the other 18 studies provided Tier IV evidence (Table 3.2).

Catterall and Pepler (2007) examined the effects of two visual arts programs serving inner-city 9-year-olds in two U.S. cities. The Inner City Arts (ICA) is an institution in downtown Los Angeles serving inner-city schools. The Center of Contemporary Arts (COCA) is a visual arts program that took place in schools in the public housing projects in St. Louis, Missouri. Both ICA and COCA featured, as per the authors, rich, sustained visual arts instruction. The study found that students who participated in the two programs made significantly greater gains than comparison students in self-efficacy beliefs and in originality (i.e., a process ability) as measured by a children’s version of a standard creativity test. The authors asserted that these effects were attributed to children’s engagement in art and to diverse positive interactions with peers and expert instructors surrounding the work.

The 18 studies that provided Tier IV evidence addressed a range of outcomes, including art learning (e.g., quality of visual art work), social-emotional learning (e.g., self-efficacy, self-esteem, self-confidence, social responsibility), process abilities (e.g., critical thinking), and attitudes (e.g., attitude toward and interest in art). No studies of visual arts interventions provided evidence aligned with Tier IV for academic achievement outcomes.

¹¹ Two of the studies examined the same intervention.

Table 3.1. Evidence Review Results: Studies of Visual Arts Interventions With Tiers II and III Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Outcomes for Which Positive Effects Were Found	Evidence Tier	Characteristics of Study Samples	Citation	Abstract Link
3	Inner City Arts (ICA): Los Angeles Center of Contemporary Arts (COCA): St Louis	<ul style="list-style-type: none"> ▶ Los Angeles, CA ▶ St. Louis, MO 	<ul style="list-style-type: none"> ▶ Teaching artists, arts specialists ▶ Sustained instruction (ICA—90 minutes twice per week for 20 weeks; COCA—60 minutes once per week for 30 weeks) 	<ul style="list-style-type: none"> ▶ Social-emotional learning: Self-efficacy ▶ Process abilities: Originality 	Tier III	<ul style="list-style-type: none"> ▶ ICA—97% of students qualifying for free- or reduced-price lunch, 97% Hispanic ▶ COCA—99% students qualifying for subsidized meals, 100% African American 	Catterall & Peppler (2007)	https://eric.ed.gov/?id=EJ779868

Table 3.2. Evidence Review Results: Studies of Visual Arts Interventions With Tier IV Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
Elementary Grades (Prekindergarten–Grade 5)					
K–5	Cross-age mentoring and artist modeling	Southern United States	<ul style="list-style-type: none"> ▶ Mentoring ▶ Modeling 	Green, Mitchell, & Taylor (2011)	https://eric.ed.gov/?id=EJ932846
PK, K	Visual literacy education	New York City, NY/New Jersey	Visual literacy interviews	Lopatovska, Hatoum, Waterstraut, Novak, & Sheer (2016)	https://www.researchgate.net/publication/308655437_Not_just_a_pretty_picture_visual_literacy_education_through_art_for_young_children
Middle and Mixed Grade Levels (up to Grade 8)					
8	Improving student attitudes and achievement in art	Midwest	<ul style="list-style-type: none"> ▶ Professional art gallery ▶ Lessons ▶ Student portfolios 	Benton (2000)	https://files.eric.ed.gov/fulltext/ED444921.pdf
K–8	Visual Thinking Strategies (VTS) curriculum	Byron, MN	<ul style="list-style-type: none"> ▶ School art teachers ▶ Curriculum designed to develop critical thinking ▶ In-school discussions about art along with museum visits 	Housen (2001)	https://files.eric.ed.gov/fulltext/ED457186.pdf
6	Peace Park Project	Western United States	<ul style="list-style-type: none"> ▶ Professional artists ▶ Youth arts organization 	Krensky (2001)	http://www.artsedsearch.org/study/going-on-beyond-zebra-a-middle-school-and-community-based-arts-organization-collaborate-for-change/
8	Online-learning component created by the teacher-researcher	Georgia	<ul style="list-style-type: none"> ▶ Online learning incorporating peer discussion groups ▶ Digital resources 	Lovin & Lambeth (2014)	https://eric.ed.gov/?id=EJ1097404
6	The sequential method of teaching perspective drawing	Washington	<ul style="list-style-type: none"> ▶ Art specialist ▶ Two 55-minute lessons on various concepts of perspective drawing. 	Snow & McLaughlin (2005)	https://eric.ed.gov/?id=EJ718125
6–8	Choice-Based Art Education (CBAE) curriculum	Philadelphia, PA	<ul style="list-style-type: none"> ▶ Student-centered inclusive curriculum that supports the needs of diverse learners ▶ 40-minute class five days a week for 36 days 	Varian (2016)	https://eric.ed.gov/?id=ED567778

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
6–8	Supplemental synchronous lessons to Florida Virtual School's M/J Orientation to Art 2D course	Florida	<ul style="list-style-type: none"> ▶ Online arts educator ▶ Web conferencing tool Blackboard Collaborate 	Werner (2011)	https://eric.ed.gov/?id=ED549416
High School and Mixed Grade Levels (up to Grade 12)					
10–12	The Artistic Impetus Model	Jersey City, NJ	<ul style="list-style-type: none"> ▶ Lessons ▶ Sensory-based exercises ▶ Reflective dialogues 	Amorino (2009)	https://eric.ed.gov/?id=EJ867993
9–12	Online art community: New Media & Social Networking in Art Education	Not specified	<ul style="list-style-type: none"> ▶ School art instructors ▶ Custom designed social networking site ▶ Curriculum designed by the teacher-researcher 	Castro (2012)	https://eric.ed.gov/?id=EJ976281
9–12	Online collaborative music creation technologies	Miami, FL	<ul style="list-style-type: none"> ▶ Computer software ▶ Online collaborative tools 	Cremata & Powell (2017)	https://eric.ed.gov/?id=EJ1138698
11	Art-making and racial climate	Philadelphia, PA	<ul style="list-style-type: none"> ▶ Classroom art teacher ▶ 8-week art classes ▶ Art projects focused on identity and racial identity 	Davis (2016)	https://files.eric.ed.gov/fulltext/ED567779.pdf
8–12	Activities based on Olivia Gude's postmodern principles	Freeland, PA	<ul style="list-style-type: none"> ▶ Classroom projects ▶ Visual art journal kit 	Ferry (2016)	https://eric.ed.gov/?id=ED567777
11 & 12	Studio art class designed to enhance creativity and thinking skills	Midwest	<ul style="list-style-type: none"> ▶ Classroom art teacher ▶ Lessons emphasizing problem solving and creativity thinking skills 	Harper (2003)	https://eric.ed.gov/?id=ED479394
9–12	Introductory high school visual arts infused with labor studies content	Bronx, NY	<ul style="list-style-type: none"> ▶ School art teacher ▶ University teacher educators ▶ Curriculum that includes artworks pertinent to labor studies 	Sosin, Bekkala, & Pepper-Sanello (2010)	https://eric.ed.gov/?id=EJ1093573
9–12	Use of computer hypertext (through a computer program called Storyspace) as a base for art education	Not specified	The Storyspace software program for creating, editing, and reading hypertext fiction.	Taylor (2000)	https://www.tandfonline.com/doi/abs/10.1080/00393541.2000.11651688
8–12	Concept-based inquiry in an arts classroom	Not specified	<ul style="list-style-type: none"> ▶ Classroom art teacher ▶ Critiques ▶ Two- or 3-day workshops 	Walker (2014)	https://eric.ed.gov/?id=EJ1040853

Note. The studies in the table are listed alphabetically by first author within each grade level.

Evidence in Studies of Dance Interventions

Four studies of dance interventions were classified as providing evidence aligned with ESSA tiers. Two studies produced evidence aligned with Tier III (Table 3.3), and the other two studies provided Tier IV evidence (Table 3.4).

One study (Lobo & Winsler, 2006) that provided Tier III evidence assessed the effects of a creative dance and movement program on the social competence of preschool children in the Head Start setting. The study reported statistically positive findings from a well-designed and well-implemented randomized controlled trial (RCT). The study therefore meets a key design requirement for Tier I. However, because the study was based on a small sample (40 students) and took place at one site, it does not meet the large-sample and multisite requirements for Tier I evidence. The study was classified as providing Tier III evidence instead.

Minton (2003) examined the effects of dance education on students' creative thinking by comparing the creative thinking ability of students taking dance classes with the ability of students not taking dance classes in six high schools. Creative thinking ability was measured using the Torrance Test of Creative Thinking (TTCT). The dance classes included a mixture of dance forms at the beginning and advanced levels, and the time students spent on dancing each week varied across classes. The study found statistically significant positive effects on two of the TTCT subscales (originality and abstractness of titles).

Two studies of dance interventions failed to meet the design requirements for Tiers I–III, but they included a research or theory-based rationale for the interventions. The studies therefore were classified as providing evidence at Tier IV. Harding and Haven (2009) examined the effect of peer coaching on transfer of skills and content into performance in dance classes in two public high schools. Leijen, Admiraal, Wildschut, and Simons (2008) investigated students' learning experience in an international¹² distance dance education program delivered in an e-learning format using a virtual learning environment platform.

¹² The study was considered eligible because a quarter of the students in the sample were from the United States.

Table 3.3. Evidence Review Results: Studies of Dance Interventions With Tiers II and III Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Outcomes for Which Positive Effects Were Found	Evidence Tier	Characteristics of Study Samples	Citation	Abstract Link
PK	Creative dance/movement program	Mid-Atlantic region	<ul style="list-style-type: none"> ▶ Small-group creative dance instruction ▶ Dance teacher 	Social-emotional learning: Social competence	Tier III	<ul style="list-style-type: none"> ▶ 40 preschool children ▶ 67% Hispanic, 16% African American, 5% Asian, 7% Arab, 5% Caucasian/Other 	Lobo & Winsler (2006)	http://www.artsedsearch.org/study/the-effects-of-a-creative-dance-and-movement-program-on-the-social-competence-of-head-start-preschoolers/
9–12	High school dance classes	Not specified	<ul style="list-style-type: none"> ▶ Dance teachers ▶ A mixture of dance forms at various levels 	Process abilities: Creative thinking	Tier III	286 students across six high schools	Minton (2003)	https://eric.ed.gov/?id=EJ824162

Table 3.4. Evidence Review Results: Studies of Dance Interventions With Tier IV Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
11 & 12	Peer coaching in dance classes	Suburban	Group or partner coaching	Harding & Haven (2009)	https://www.tandfonline.com/doi/abs/10.1080/15290824.2009.10387385
Not stated	E-Learning and dance (use of Claroline Learning Environment)	International, including 11 U.S. students	<ul style="list-style-type: none"> ▶ Dance teachers ▶ Dance program with four e-courses per student 	Leijen, Admiraal, Wildschut, & Simons (2008)	https://www.tandfonline.com/doi/abs/10.1080/14647890802087951

Evidence in Studies of Drama Interventions

Our review found eight studies of drama interventions that provided evidence at Tiers III and IV. Two studies produced evidence aligned with Tier III (Table 3.5), and six studies provided Tier IV evidence (Table 3.6).

Greene, Erickson, Watson, and Beck (2017) and Kisiel et al. (2006) both provided Tier III evidence. Greene et al. (2017) reported the results of five random assignment experiments in which classes were assigned by lottery to attend a live theater performance or to watch a movie version of the same story. The intervention examined in the study did not include anything beyond the opportunity to see live theater. The study found that students in the classes assigned to see live theater scored significantly higher on measures of tolerance and content knowledge (e.g., stronger command of the plot and vocabulary of those plays) than students in classes assigned to watch a movie. Kisiel et al. (2006) evaluated the impact of Urban Improv, a theater-based youth violence prevention program developed for inner-city youth, on a range of social-emotional outcomes (e.g., aggressive behaviors, prosocial behaviors, cooperation, and self-control) using a quasi-experimental design. The study found statistically significant positive effects of Urban Improv on measures of students' cooperation, self-control, hyperactivity, and internalizing behavior, based on ratings provided by teachers.

Six studies of drama interventions provided rationales needed for Tier IV evidence (Table 3.6). Four studies provided rationales linking specific drama interventions to various social-emotional learning outcomes such as self-concept, problem behavior, social skills, self-efficacy, and anger inhibition (Bernstein, Ablow, Maloney, & Nigg, 2014; Freeman, Sullivan, & Fulton, 2003; Goldstein & Winner, 2012; Holloway & LeCompte, 2001). One report provided a rationale linking a drama intervention (videotaping as a means of self-monitoring) with students' theater performance (Lan & Morgan, 2003). The other report provided a rationale linking a drama intervention (improvisational learning activities in drama classes) with students' critical thinking skills (Chou & Shih, 2010).

Table 3.5. Evidence Review Results: Studies of Drama Interventions With Tier III Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Outcomes for Which Positive Effects Were Found	Evidence Tier	Characteristics of Study Samples	Citation	Abstract Link
4–12	Live theater performance	Arkansas	Field trip to see live theater performance	<ul style="list-style-type: none"> ▶ Art learning: Content knowledge ▶ Social-emotional learning: Tolerance 	Tier III	94 classes, approximately 1,500 students	Greene, Erickson, Watson, & Beck (2017)	http://journals.sagepub.com/doi/abs/10.3102/0013189X18761034
4	Urban Improv (UI)	Boston, MA	<ul style="list-style-type: none"> ▶ Interactive program design using structured theater improvisation ▶ UI intervention staff ▶ 9-week developmentally appropriate curriculum 	<ul style="list-style-type: none"> ▶ Social-emotional learning: Self-control, cooperation, hyperactivity, internalizing behavior 	Tier III	<ul style="list-style-type: none"> ▶ 140 students from eight fourth-grade classrooms ▶ 44.5% African American, 27.7% Hispanic, 13.1% biracial, 5.8% Asian, 5.8% Other, 2.2% Caucasian, 0.7% Native American 	Kisiel, Blaustein, Spinazzola, Schmidt, Zucker, & van der Kolk (2006)	https://www.tandfonline.com/doi/abs/10.1300/J202v05n02_03

Table 3.6. Evidence Review Results: Studies of Drama Interventions With Tier IV Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
Elementary Grades (Prekindergarten–Grade 5)					
3 & 4	Creative drama	Rural	<ul style="list-style-type: none"> ▶ Creative drama instruction ▶ Planning by drama and music teachers 	Freeman, Sullivan, & Fulton (2003)	https://www.tandfonline.com/doi/abs/10.1080/00220670309598801
Middle Grades (Grades 6–8)					
6	Dramatic improvisational activity	Cambridge, MA	<ul style="list-style-type: none"> ▶ Lessons ▶ Class discussion 	Chou & Shih (2010)	https://eric.ed.gov/?id=EJ1093600
6–8	Arts Focus	Not specified	<ul style="list-style-type: none"> ▶ Intensive art instruction (90-minute daily) ▶ Innovative teaching 	Holloway & LeCompte (2001)	https://eric.ed.gov/?id=EJ631870
High School Grades (Grades 9–12)					
9 & 10	PlayWrite	Midsized Urban Alternative high schools	<ul style="list-style-type: none"> ▶ Workshops ▶ Playwriting with coaches ▶ Live performance 	Bernstein, Ablow, Maloney, & Nigg (2014)	https://www.tandfonline.com/doi/abs/10.1080/15401383.2014.902342
High school freshmen (Grade 9)	Acting training	Suburban and large urban high schools for the arts	<ul style="list-style-type: none"> ▶ High-intensity acting training (9 hours per week) ▶ Productions 	Goldstein & Winner (2012)	https://www.tandfonline.com/doi/abs/10.1080/15248372.2011.573514
High school	Self-monitoring	Urban	<ul style="list-style-type: none"> ▶ Play development ▶ Videotaping ▶ Self-monitoring 	Lan & Morgan (2003)	https://www.tandfonline.com/doi/abs/10.1080/00220970309602070?journalCode=vjxe20

Note. The studies in the table are listed alphabetically by first author within each grade level.

Evidence in Studies of Music Interventions

The majority of studies that were relevant to this evidence review involve music interventions. Our review found 50 studies of music interventions that provided evidence at Tiers II–IV. One study produced evidence aligned with Tier II, eight studies produced evidence aligned with Tiers III (Table 3.7), and 41 studies provided Tier IV evidence (Table 3.8).

The study that provided Tier II evidence (Helmrich, 2010) compared the algebra achievement of students receiving formal instrumental or choral music instruction during middle school (the instrumental group or the choral group) with those who received neither of those modes of musical instruction (the comparison group) in six school districts in Maryland. The study found that students in the choral group outperformed the comparison group, and the difference was statistically significant. This finding was considered to provide Tier II evidence because the two groups were shown to be equivalent at baseline.¹³

Five studies provided Tier III evidence about the impact of music interventions on students' academic achievement outcomes (Gromko, 2005; Miksza, 2010; Piro & Ortiz, 2009; Sharpe, 2013; Southgate & Roscigno, 2009).¹⁴ Three studies (Gromko, 2005; Piro & Ortiz, 2009; Sharpe, 2013) used quasi-experimental designs and two (Miksza, 2010; Southgate & Roscigno, 2009) were correlational studies. Gromko (2005) found that kindergarten children who received 4 months of music instruction showed significantly greater gains in their phoneme segmentation fluency than children who did not receive music instruction. Miksza (2010) examined the relationships between participating in high school music ensembles and students' mathematics achievement, community ethic, and commitment to school. The study found that students participating in band, chorus, or orchestra had significantly higher math achievement scores and significantly higher levels of community ethic and commitment to school than students who did not participate in those activities. Piro and Ortiz (2009) found that students who participated in a scaffolded music instruction program had significantly higher vocabulary and verbal sequencing scores than students in the comparison group.¹⁵ Sharpe (2013) found that Grade 6 students who received music instruction through band, chorus, or orchestra during the school year scored significantly higher in mathematics achievement tests than students who did not receive music instruction. Using nationally representative samples, Southgate and Roscigno (2009) found that in-school music participation¹⁶ was positively associated with mathematics and reading achievement among children and adolescents.

Among the other three studies that provided Tier III evidence, one study focused on art learning outcomes (Johnson, 2011), one study focused on process abilities (Rauscher & Zupan, 2000),

¹³ The study also found statistically significant positive group differences in algebra achievement based on other comparisons (e.g., comparison between the instrumental group and the neither-instruction group, and subgroup comparisons). However, those findings did not meet Tier II criteria because the groups compared were not shown to be equivalent at baseline. Those findings were considered providing Tier III evidence because the author included statistical controls (including the pretest scores) in the analyses.

¹⁴ One study (Miksza, 2010) also provided Tier III evidence on social-emotional learning outcomes.

¹⁵ This intervention was also described in another study report (Piro, 2009), which was rated as providing no tier-aligned evidence.

¹⁶ The study also examined music involvement outside of school and parental involvement in the form of concert attendance, which is out of the scope of this review. Our review focused on only in-school music involvement.

and the other study focused on social-emotional learning outcomes (Ritblatt, Longstreth, Hokoda, Cannon, & Weston, 2013). Johnson (2011) examined the effects of including critical thinking opportunities in music instruction (compared with activity-based music instruction that did not include opportunities for critical thinking) on music listening skills of Grade 5 students. Statistically significant differences were found between the treatment and comparison groups on the music listening and critical thinking measures. Ritblatt et al. (2013) assessed the effects of the Circle of Education, a music-based school-readiness program, on preschool children's social skills and school-readiness skills. The study found that the improvement in social skills (total score) and specifically on the social cooperation, social interaction, and social independence scales were significantly higher for the music group than for the comparison group. Rauscher and Zupan (2000) reported the results of a field experiment in which kindergarteners were assigned to receive music instruction featuring the keyboard or to receive no music instruction. The study found that the keyboard group scored significantly higher than the no-music group on spatial-temporal tasks after 4 months of lessons, and that the difference was significantly greater after 8 months of lessons.

Forty-one studies of music interventions did not meet the criteria for Tiers I–III evidence but did provide rationales needed for Tier IV evidence (Table 3.8). These studies provided rationales linking music education interventions to multiple outcome domains examined in this review. For example, 21 studies provided rationales linking music education interventions to art learning outcomes, 11 studies linked music education to social-emotional learning outcomes, eight studies linked music education to student achievement outcomes, eight studies linked music education to process abilities, eight studies linked music education to attitudes, and four studies linked music education to other outcomes within the scope of this review (e.g., later life involvement in music, outlook for life, and instructional strategies).

Table 3.7. Evidence Review Results: Studies of Music Interventions With Tiers II and III Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Outcomes for Which Positive Effects Were Found	Evidence Tier	Characteristics of Study Samples	Citation	Abstract Link
Elementary Grades (Prekindergarten–Grade 5)								
K	Music instruction	Midwest United States	<ul style="list-style-type: none"> ▶ Classroom teachers ▶ University music method students ▶ 4 months of music instruction 	Academic achievement: Phoneme - segmentation fluency	Tier III	103 kindergarteners at two elementary schools	Gromko (2005)	http://journals.sagepub.com/doi/abs/10.1177/002242940505300302
5	Critical Thinking Instruction (CTI)	Southwest United States	Music instruction using open-ended, higher order cognitive questions and improvisation exercises	Art learning: Music listening skills	Tier III	81 fifth-grade students	Johnson (2011)	http://journals.sagepub.com/doi/abs/10.1177/0022429411415860
2	Keyboard instruction based on the Music and the Brain project curriculum	New York City	<ul style="list-style-type: none"> ▶ Music teacher ▶ The Music and the Brain project curriculum ▶ A specially equipped music laboratory 	Academic achievement: Vocabulary, verbal sequencing	Tier III	<ul style="list-style-type: none"> ▶ 103 second-grade students ▶ 54% non-White ▶ 53% receive free lunch 	Piro & Ortiz (2009)	http://journals.sagepub.com/doi/10.1177/0305735608097248
K	Classroom keyboard instruction	Midwest United States	<ul style="list-style-type: none"> ▶ Music specialist ▶ 20-min lessons twice per week in groups of approximately 10 children 	Process abilities: Spatial-temporal reasoning	Tier III	<ul style="list-style-type: none"> ▶ 62 middle-income kindergarten children ▶ 36 boys and 26 girls ▶ Mixed ethnicity 	Rauscher & Zupan (2000)	http://psycnet.apa.org/record/2001-03185-003
PK	Circle of Education music-based school-readiness program	Southwest United States	<ul style="list-style-type: none"> ▶ 20 songs created based on the Desired Results Developmental Profile focusing on social-emotional skills ▶ Child care instructors 	Social-emotional learning: Social cooperation, social interaction, social independence	Tier III	102 preschool students	Ritblatt, Longstreth Hokoda, Cannon, & Weston (2013)	https://www.tandfonline.com/doi/abs/10.1080/02568543.2013.796333?scroll=top&needAccess=true&journalCode=uirc20

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Outcomes for Which Positive Effects Were Found	Evidence Tier	Characteristics of Study Samples	Citation	Abstract Link
Middle Grades (Grade 6–8)								
6–8	Middle school music programs	Maryland	Instrumental or choral instruction all 3 years during middle school	Academic achievement: Algebra achievement	Tier II	6,026 students across six school districts	Helmrich (2010)	http://journals.sagepub.com/doi/10.1177/0743558410366594
6	Music Instruction and Academic Achievement in Mathematics	Atlanta, GA	Music instruction through band, chorus, or orchestra	Academic achievement: Math achievement	Tier III	109 sixth-grade students	Sharpe (2013)	https://eric.ed.gov/?id=ED563908
High School and Mixed Grade Levels (up to Grade 12)								
10	School music ensemble participation	National	High school music ensembles	<ul style="list-style-type: none"> ▶ Academic achievement: Math achievement ▶ Social-emotional learning: Community ethic, commitment to school 	Tier III	12,160 students across 603 high schools	Miksza (2010)	http://www.jstor.org/stable/41110431
K–1, 8–12	Music involvement and academic achievement	National	School music programs	Academic achievement: Reading and math achievement	Tier III	<ul style="list-style-type: none"> ▶ 4,376 students in the ECLS-K sample ▶ 7,781 students in the NELS:88 sample 	Southgate & Roscigno (2009)	http://psycnet.apa.org/record/2009-00725-001

Note. The studies in the table are listed alphabetically by first author within each grade level. ECLS-K refers to the Early Childhood Longitudinal Study, Kindergarten Class data; NELS:88 refers to the National Education Longitudinal Study of 1988 data. Both data sets were collected and maintained by the National Center for Education Statistics.

Table 3.8. Evidence Review Results: Studies of Music Interventions With Tier IV Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
Elementary Grades (Prekindergarten–Grade 5)					
5	Multicultural instructional approaches	4 Urban schools	<ul style="list-style-type: none"> ▶ Music concept with formal elements of music instruction ▶ Sociocultural music instruction 	Abril (2006)	http://journals.sagepub.com/doi/pdf/10.1177/0255761406063103
PreK	Music therapy	Not reported	Using instrument play, body movement, listening, and vocalizations in music therapy	Barnes (2010)	https://eric.ed.gov/?id=ED527813
PreK	Pre-K music instruction	Southeast United States	<ul style="list-style-type: none"> ▶ Curriculum based on the National Music Standards for prekindergarten ▶ Music instruction 	Bowen (2010)	https://eric.ed.gov/?id=ED519022
3	Mathematics-related music instruction	Southeast GA	<ul style="list-style-type: none"> ▶ Music instruction ▶ Classroom music teacher 	Brock & Lambeth (2013)	https://eric.ed.gov/?id=ED565237
2–4	Elementary music activities and exercises for creativity	Midwest	<ul style="list-style-type: none"> ▶ Various activities: listening homework exercise, in class activity ▶ Teacher played music and demonstrated skills 	Coulson & Burke (2013)	http://journals.sagepub.com/doi/abs/10.1177/0255761413495760
K–5	Use of a recorded aural model	Not specified	<ul style="list-style-type: none"> ▶ School music teacher ▶ Recorded aural model 	Frewen (2010)	https://eric.ed.gov/?id=EJ877133
K & 1	Pedagogical approaches and present/absence of text in music instruction	Suburban, CT	<ul style="list-style-type: none"> ▶ Traditional folk songs ▶ Four treatment conditions varying in the way songs were taught ▶ Varying holistic teaching and text conditions (i.e., presence or absence of text) 	Gault (2002)	https://www.jstor.org/stable/40319126?seq=1#page_scan_tab_contents
5	Self-monitoring: teacher/researcher designed action research	Rural, Midwest	Use of student self-monitoring checklist and other supports toward developing and using higher order thinking skills	Hayes (2002)	https://files.eric.ed.gov/fulltext/ED471585.pdf

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
K	Intensive music program	New England	Increased time in general music classes (two to seven times more than control students)	Hogan, Cordes, Holochwost, Ryu, Diamond, & Winner (2018)	https://doi.org/10.1016/j.ecresq.2017.12.004
K-2; 4 & 5	Musical ensembles in residence	Southwestern United States	Enhancement to regular music program with ensemble interaction once a week	Johnson & Davis (2016)	https://openmusiclibrary.org/article/172629/
1-3	Harmony Project	Los Angeles, CA	<ul style="list-style-type: none"> ▶ Community music program ▶ Curriculum ▶ Music instruction 	Kraus, Hornickel, Strait, Slater, & Thompson (2014)	https://www.ncbi.nlm.nih.gov/pubmed/25566109
PreK	Music training in Head Start setting	Not specified	<ul style="list-style-type: none"> ▶ Music training ▶ Music activities ▶ Classroom teachers 	Neville et al. (2008)	http://lup.lub.lu.se/record/3364201
3	SingingCoach software	Midwest to Northwest states	Secondary data analysis of music classrooms participating in the Great American Singing Challenge	Paney & Kay (2015)	http://journals.sagepub.com/doi/full/10.1177/8755123314548047
4 & 5	Instruction and melody and improvisation	Georgia	<ul style="list-style-type: none"> ▶ School music teacher ▶ Instruction 	Parisi (2004)	https://eric.ed.gov/?id=EJ845396
K	Live Music Between the Lions TV	Florida	<ul style="list-style-type: none"> ▶ Music groups ▶ Television program 	Register (2004)	https://academic.oup.com/jmt/article-abstract/41/1/2/1035029?redirectedFrom=fulltext
2	Solfège syllables and relating tonal patterns	Northeast	Sight singing instruction with related and nonrelated (syllable) songs	Reifinger (2012)	https://www.jstor.org/stable/41348850?seq=1#page_scan_tab_contents
1	Teacher feedback and modeling	Pennsylvania	School music teacher	Rutkowski & Miller (2003)	https://www.jstor.org/stable/40319169?seq=1#page_scan_tab_contents
PreK & K	Early childhood music class	Northeastern United States	Music class collaboration	St. John (2006)	http://journals.sagepub.com/doi/10.1177/0305735606061854
PreK	Kindermusik	<ul style="list-style-type: none"> ▶ MidAtlantic ▶ Suburban or metropolitan area 	Music-based curriculum	Winsler, Ducenne, & Koury (2011)	https://www.tandfonline.com/doi/abs/10.1080/10409280903585739

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
Middle and Mixed Grade Levels (up to Grade 8)					
6–8	School-loaner instrument program	Urban	Providing instruments to students who could not afford to own their own	Ester & Turner (2009)	https://eric.ed.gov/?id=EJ864262
5–8	Self-evaluation instruction	Metropolitan mid-Atlantic	Instruction in self-evaluation	Hewitt (2011)	http://journals.sagepub.com/doi/10.1177/0022429410391541
7	Peer Assisted Learning (PAL) structures	Urban/suburban	Symmetrical and asymmetrical (PAL) groups	Johnson (2017)	http://journals.sagepub.com/doi/abs/10.1177/0022429417712486
K–8	Music education programs	South, East Coast, Midwest, West Coast	<ul style="list-style-type: none"> ▶ Music programs ▶ Music instruction 	Johnson & Memmott (2006)	https://eric.ed.gov/?id=EJ773246
6–8	Peer Connectedness Study	Suburban, upper Midwest	School-based instrumental music instruction and band-related activities	Rawlings & Stoddard (2017)	http://journals.sagepub.com/doi/abs/10.1177/1321103X17703575
8	Teaching composition in middle school band	Not specified	<ul style="list-style-type: none"> ▶ Teaching composition along with performance and listening to wind and percussion players ▶ Use of SCAMPER strategies in some activities 	Riley (2006)	http://journals.sagepub.com/doi/abs/10.1177/87551233060250010104
6	Music education and mentoring	Midwest	General instruction alternates with performance-oriented and product-oriented	Shields (2001)	http://journals.sagepub.com/doi/pdf/10.2307/3345712
5 & 6	Melodic error detection	Not reported	<ul style="list-style-type: none"> ▶ Listening to themselves while playing ▶ Listening to recordings ▶ Familiar melodies 	Thornton (2008)	http://journals.sagepub.com/doi/pdf/10.1177/8755123308317612
7	A curriculum to increase music achievement using technology, self-assessment, and positive feedback	Urban	<ul style="list-style-type: none"> ▶ MIDI capable workstations ▶ Music Shop software ▶ Portfolio development ▶ Pentachord unit 	Vega (2001)	https://files.eric.ed.gov/fulltext/ED456086.pdf

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
Middle school	Little Kids Rock	Dallas, TX	<ul style="list-style-type: none"> ▶ Free instruments ▶ Teacher training ▶ Curricular materials ▶ Music instruction 	Thomas (2016)	https://www.arts.gov/sites/default/files/Research-Art-Works-MississippiState.pdf
High School and Mixed Grade Levels (up to Grade 12)					
Adolescents	Music mentorship program	Not identified	Music with mentorship of students with developmental disabilities	Darrow, Novak, Swedberg, Horton, & Rice (2009)	https://eric.ed.gov/?id=EJ912418
High school	Music class participation	National	School music offerings	Elpus (2013)	http://journals.sagepub.com/doi/abs/10.1177/0022429413485601?journalCode=jrma
9–12	Comprehensive Musician (CMP)	Illinois	<ul style="list-style-type: none"> ▶ School band directors ▶ Lesson plans ▶ Unit study-technology approach 	Gustafson-Hinds (2010)	http://irl.umsl.edu/dissertation/486
9–12	Music history and theory instruction	Not reported	Music history and theory lessons	Hendricks (2010)	http://www.jstor.org/stable/27861483
9–12	Sight-singing instruction and specific pitch skills	Central Texas	<ul style="list-style-type: none"> ▶ School music teacher ▶ Solfege drill 	Henry (2004)	https://eric.ed.gov/?id=EJ708660
7–9	Modeling, self-evaluation, and self-listening	Southwest	Varied combinations of modeling, self-listening, and self-evaluation before and during instrumentalists' music performance	Hewitt (2001)	http://journals.sagepub.com/doi/10.2307/3345614
9–12	The Lakewood Project	Midwest	<ul style="list-style-type: none"> ▶ Use of classical instruments and rock band instruments ▶ Focus on informal learning integrated with traditional teaching 	Koops, Hankins, Scalise, & Schatt (2014)	http://journals.sagepub.com/doi/full/10.1177/1321103X14537432
Mixed high school	Music composition instruction	Midwest	<ul style="list-style-type: none"> ▶ Composition instruction ▶ Teacher orientation 	Menard (2015)	http://journals.sagepub.com/doi/abs/10.1177/0022429415574310

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
6–12	Using recorded models in ensemble rehearsals	<ul style="list-style-type: none"> ▶ Rural, urban, and suburban ▶ Pacific Northwest 	Recorded models of professional recordings used to prepare for selected pieces	Morrison, Montemayor, & Wiltshire (2004)	http://journals.sagepub.com/doi/10.2307/3345434
6–12	Mariachi Education Program	Nevada, Georgia, Tennessee, California	<ul style="list-style-type: none"> ▶ Curriculum ▶ Teacher PD ▶ Music instruction 	Neel (2017)	https://eric.ed.gov/?id=EJ1154274
High school	High school music classes	Chicago, Illinois	<ul style="list-style-type: none"> ▶ Two years of high school music classes 	Tierney, Krizman, Skoe, Johnston, & Kraus (2013)	http://www.frontiersin.org/journal/10.3389/fpsyg.2013.00855/abstract
Adolescents	Weill Music Institute's strength-based music program in juvenile justice settings	<ul style="list-style-type: none"> ▶ Urban ▶ Northeast 	<ul style="list-style-type: none"> ▶ Twelve sessions over 2 weeks ▶ Music focus 	Wolf & Holochwost (2014)	http://wolfbrown.com/images/books_reports/documents/ourvoicescount.pdf

Note. The studies in the table are listed alphabetically by first author within each grade level.

Evidence in Studies of Interventions That Involve Multiple Types of Art

Seven studies provided evidence aligned with ESSA tiers on interventions that involved more than one type of art. Two studies provided evidence at Tier II, two studies provided evidence at Tier III, and three studies provided evidence at Tier IV.

Both interventions supported by Tier II evidence took place in New York City (Table 3.9). Chen, Lui, Andrade, Valle, and Mir (2017) examined the effect of professional development in the use of criteria-referenced formative arts assessment on students' achievement in arts. The study was part of the Arts Achieve project, funded by a U.S. Department of Education Investing in Innovation Fund grant and an Arts in Education Model Development and Dissemination grant. The treatment group consisted of students taught by teachers in all four major art disciplines (i.e., music, visual arts, drama, and dance) who received the professional development. The control group consisted of students in classes instructed by teachers who did not receive the professional development. The study showed that professional development in the use of the criteria-referenced formative assessment in arts had a statistically significant positive effect on students' arts achievement, as measured by discipline-specific, performance-based assessments.

Horowitz (2016) examined the impact of the Everyday Arts for Special Education (EASE) program—also funded through an Investing in Innovation Fund development grant to New York City's special education district (District 75)—on special education students' academic achievement and social-emotional learning. The participating EASE students were in grades 2–5 in 10 schools. EASE teachers participated in professional development and received in-school support on strategies across multiple arts disciplines (i.e., music, dance, visual arts, and theater). The study found statistically significant effects on social-emotional learning outcomes and on reading achievement. However, only the findings on social-emotional outcomes met the criteria for Tier II evidence; the findings on reading achievement were based on a sample that is smaller ($N = 157$ students) than the threshold for Tier II evidence and were therefore classified as providing Tier III evidence.

Like Chen et al.'s (2017) study, Mastrorilli, Harnett, and Zhu (2014) also examined the impact of professional development in the use of criteria-referenced formative assessment in arts as part of the Arts Achievement project, but with a slightly different sample. The study provided Tier III evidence on students' arts achievement. Neddeau (2013) provided Tier III evidence on the relationship between teachers' reported curriculum emphasis in arts (music, art, dance or creative movement, and theater or creative dramatics) and students' reading achievement. Teachers' curriculum emphasis in arts was measured by instructional minutes per week. The study found a statistically significant positive relationship between teachers' curriculum emphasis in arts and students' reading achievement in Grade 5.

Three studies of interventions involving multiple types of art did not meet the criteria for evidence at Tiers I–III but did provide rationales needed for Tier IV evidence (Table 3.10). Abedin (2010) investigated the learning experience of adolescents with learning disabilities in their

music and drama classes during an 8-week curriculum unit in a public charter school with a quality arts program. Lovett (2014) examined the relationship between fine arts courses (dance, music, theatre arts, and visual arts) offered in a school and the graduation rate of the school among Georgia's public high schools. Rapp-Paglicci, Stewart, and Rowe (2012) evaluated the impact of the Prodigy Cultural Arts Program, an early prevention program for adjudicated youth, on students' mental health, academic achievement, and family functioning.

Table 3.9. Evidence Review Results: Studies of Interventions Involving Multiple Art Types With Tiers II and III Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Outcomes for Which Positive Effects Were Found	Evidence Tier	Characteristics of Study Samples	Citation	Abstract Link
K-12	PD on criteria-referenced formative assessment in the arts as part of the Arts Achieve project	New York City, NY	<ul style="list-style-type: none"> ▶ Art teachers ▶ Teacher PD ▶ Teacher PLCs 	Art learning: Arts assessment scores	Tier II	<ul style="list-style-type: none"> ▶ 5,640 students across 75 schools ▶ 72% minority 	Chen, Lui, Andrade, Valle, & Mir (2017)	https://eric.ed.gov/?id=EJ1152545
2-5	Everyday Arts for Special Education (EASE)	New York City, NY	<ul style="list-style-type: none"> ▶ Teacher PD workshops ▶ On-site PD ▶ Collaborative classroom modeling ▶ Classroom instruction 	<ul style="list-style-type: none"> ▶ Academic achievement: Reading and math achievement ▶ Social-emotional learning: Functional skills 	Tier II	<ul style="list-style-type: none"> ▶ Special education students ▶ Academic achievement analysis included 157 Grade 4 students within 10 schools ▶ Analysis of social-emotional behavior included 759 students within 10 schools 	Horowitz (2016)	https://eric.ed.gov/?id=ED573547
K-12	PD on criteria-referenced formative assessment in the arts as part of the Arts Achieve project	New York City, NY	<ul style="list-style-type: none"> ▶ Teaching artists ▶ Teacher PD ▶ Teacher PLCs 	Art learning: Arts assessment scores	Tier III	<ul style="list-style-type: none"> ▶ 4,066 students ▶ 85% minority ▶ 11% English learners ▶ 13% special education students ▶ 85% receiving free or reduced-price lunch 	Mastrorilli, Harnett, & Zhu (2014)	http://www.artsedse.arch.org/study/h-the-arts-101-1-24/
1, 3, 5	Curricular emphases and academic achievement	United States	<ul style="list-style-type: none"> ▶ Classroom teachers ▶ Curricular emphasis in arts 	Academic achievement: Reading achievement	Tier III	Students in Grade 5 ECLS-K sample	Neddeau (2013)	https://repository.usfca.edu/cgi/viewcontent.cgi?article=1090&context=diss

Note. The studies in the table are listed alphabetically by first author. PD = professional development. PLC = professional learning community.

Table 3.10. Evidence Review Results: Studies of Interventions Involving Multiple Art Types With Tier IV Evidence

Grade Level	Name of Intervention	Geographic Location	Intervention Components	Citation	Abstract Link
7 & 8	Art-based education for adolescents with learning disabilities	<ul style="list-style-type: none"> ▶ Urban ▶ Public charter school 	Art, music, and drama classes supporting curriculum unit on Middle Eastern and African arts	Abedin (2010)	https://drum.lib.umd.edu/handle/1903/10505
9–12	Fine arts courses	Georgia	Dance, music, theatre arts, and visual arts courses	Lovett (2014)	https://eric.ed.gov/?id=ED557020
Adolescents, ages 10–18	Prodigy Cultural Arts based on Positive Youth Development Model	Urban	<ul style="list-style-type: none"> ▶ Classes in performing arts taught by master artists from the community ▶ Three hours per week for 8 weeks 	Rapp-Paglicci, Stewart, & Rowe (2012)	https://www.tandfonline.com/doi/abs/10.1080/15433714.2011.581532

Summary

This chapter presents findings for the first research question that motivated this review: Are there research studies on arts education interventions that meet the criteria for evidence as specified in ESSA? These findings are summarized by type of art.

For visual arts interventions:

- ▶ One study provided Tier III evidence suggesting that the students who participated in two inner-city programs that featured rich, sustained visual arts made significant greater gains than comparison students in social-emotional learning outcomes and process abilities.
- ▶ Eighteen studies provided Tier IV evidence by including a research- or theory-based rationale for how visual arts interventions should impact students' art learning, social-emotional learning, process abilities, and attitudes toward arts.

For dance interventions:

- ▶ Two studies produced Tier III evidence on the positive effects of dance interventions on the creative thinking abilities of high school students and on the social competence of preschool children.
- ▶ Two other studies provided Tier IV evidence supporting dance interventions by presenting research- or theory-based rationales for such interventions.

For drama interventions:

- ▶ Two studies provided Tier III evidence supporting drama interventions. One study found that students who were randomly assigned to see live theater scored significantly higher on measures of tolerance and art content knowledge than students in the control group. The other study provided evidence on the positive effects of Urban Improv, a theater-based youth violence prevention developed for inner-city youth, on a range of social-emotional outcomes.
- ▶ Six studies provided Tier IV evidence supporting drama interventions by presenting research- or theory-based rationales for such interventions.

For music interventions:

- ▶ One study provided Tier II evidence on the effect of formal instrumental or choral music instruction during middle school on students' algebra achievement.
- ▶ Eight studies provided Tier III evidence. Five studies examined the impact of music interventions on students' academic achievement outcomes, one study focused on art learning outcomes, one study focused on social emotional learning outcomes, and the other study focused on process abilities.
- ▶ Forty-one studies provided Tier IV evidence supporting music interventions by presenting research- or theory-based rationales linking music interventions to all outcome domains examined in this review.

For interventions that involved multiple types of art:

- ▶ Two studies provided Tier II evidence supporting the effectiveness of two such interventions implemented in New York City. One study examined the effectiveness of professional development in the use of criteria-referenced formative assessment in arts as part of the Arts Achieve project on students' achievement in arts. The other study examined the impact of the Everyday Arts for Special Education program on special education students' academic achievement and social-emotional learning.
- ▶ Two studies provided Tier III evidence. One study assessed the effect of professional development in the use of criteria-referenced formative assessment in arts on students' achievement in arts. The other study examined the relationships between teachers' reported curriculum emphasis in arts (music, art, dance or creative movement, and theater or creative dramatics) and students' reading achievement.
- ▶ Three studies provided Tier IV evidence supporting arts education interventions that involved multiple types of art. That is, these studies include a research- or theory-based rationale for how such interventions should impact student outcomes.

Chapter 4. How Large Are the Effects of Arts Education Interventions on Student Outcomes?

Chapter Highlights

- ▶ The average effect found in well-designed and well-implemented studies (i.e., those having research designs capable of producing evidence at Tiers I–III) was statistically significant and moderate in magnitude.
- ▶ The effects of arts education interventions varied by art type and outcome domain.

While Chapter 3 focuses on the levels of evidence supporting arts education interventions, this chapter focuses on the size of the effects that educators might expect if they adopt an arts education intervention. Specifically, this chapter addresses the second research question for this evidence review:

Research question 2: How large are the effects of arts education interventions on student outcomes?

Although the classification of interventions based on ESSA evidence tiers is useful for policymakers to ensure that educational agencies implement interventions that are supported by research, this approach has several conceptual limitations. First, the classification is biased toward statistically significant findings. As a hypothetical example, a study can make 100 statistical comparisons between an arts education group and a comparison group, and 99 of the comparisons could show no statistically significant differences. However, the one statistically significant effect would be sufficient for the intervention to be classified as providing evidence at Tiers I–III, provided that other studies on the same intervention did not show a statistically significant negative effect.

Second, the classification of interventions based on the ESSA evidence tiers is biased toward large studies and multisite studies. In one way, this bias makes sense in that larger studies can generate estimates of impact that are more accurate than smaller studies. However, the numbers of sites and individuals involved in large-scale studies may stretch the abilities of arts education intervention developers to monitor the quality of implementation and address deviations from the intended program model. In reviewing the research evidence, some allowance should be given to small studies that could have implemented arts education interventions with greater fidelity.¹⁷

Third, for school and district administrators who are considering investing in an arts education intervention, the rating of the intervention in terms of the ESSA evidence tiers fails to provide them with an idea of the magnitude of improvement in student outcomes that can be expected when adopting the intervention. Such information requires an examination of all the effects

¹⁷ In this review, we did not examine the fidelity of implementation of the interventions that were the focus of the studies reviewed. This is consistent with the convention used in the WWC review process.

produced from well-designed and well-implemented studies of the intervention, regardless of whether they produced statistically significant findings.

To address these limitations and better understand the expected effects of arts education interventions, we examined findings about the effects of arts education interventions from all 20 studies that meet the design-related criteria¹⁸ for Tiers I–III and contain effect estimates that could be converted to an effect size¹⁹ (or provided enough information for reviewers to calculate effect sizes and standard errors), even if their results did not meet the ESSA evidence standards. We examined, for example, findings that were not statistically significant, findings from studies with a sample size smaller than 350 students, and findings from studies conducted in a single site.

For each of the 20 studies, we examined the magnitude of the differences in student outcomes between the arts education intervention group and the comparison group using meta-analysis procedures. We began by converting group differences in each study to the Hedges' *g* statistic, which is an expression of the magnitude of the difference between two groups in terms of the standard deviation of the outcome measure. Conversion to the Hedges' *g* allows us to compare the effects observed based on different outcome measures in different studies using a single, common metric. Box 4.1 provides a brief explanation of our meta-analysis procedures. See Appendix A for more details on the methods used.

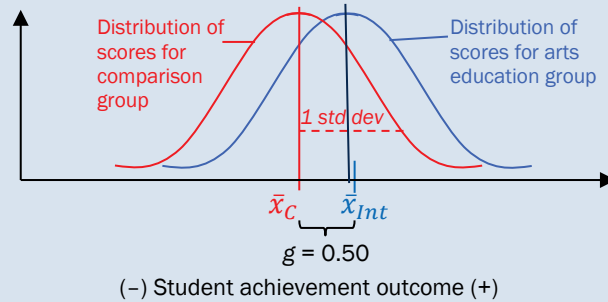
Box 4.1. Meta-Analysis Explained

A meta-analysis is a set of quantitative procedures used to statistically combine the effects from multiple studies (Cooper, 2010). It can help researchers and policymakers better understand the magnitude of the effect that an intervention can have as well as the variability of effects found with different types of samples or settings. The meta-analysis that produced the findings presented in this chapter involved the following six steps:

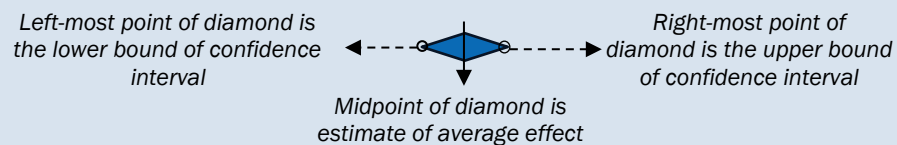
1. **Recording details about each study.** For each study that met the design criteria for ESSA's Tiers I–III (regardless of the sample size or the number of sites of the study), the review team recorded the characteristics of the intervention, research design, study setting (e.g., urban vs. suburban vs. rural), sample characteristics (e.g., student demographic information), sample size, outcome measures, and statistics about the effect of the intervention on each relevant outcome.
2. **Converting effect statistics into a common effect size metric.** The statistics reported in studies indicate the differences between students exposed to the intervention and students not exposed to the intervention on some outcome measure (or in some cases, the relationship between students' exposure to arts education and an outcome). These statistics were converted into a common effect size metric—Hedges' *g*, which represents a standardized mean difference in terms of the pooled standard deviation of the outcome measure. (The figure below illustrates an effect size of 0.50 in terms of Hedges' *g*.)

¹⁸ A study that meets the design-related criteria should be a well-designed and well-implemented RCT or quasi-experimental study (e.g., based on WWC standards) or a correlational study with statistical controls. The study may not have a large sample or a multisite sample as required for Tier I and Tier II evidence in the Non-Regulatory Guidance (U.S. Department of Education, 2016).

¹⁹ The study needs to provide effect size estimates and standard errors or provide statistics that could be used to calculate effect sizes and standard errors.



3. **Determining the standard error for each effect size.** Each effect size was estimated based on a particular sample, and there is some degree of uncertainty about whether the effect size estimate reflects the *true effect* for the population. The amount of uncertainty, measured by the *standard error*, reflects to a large extent the sample size of the study. All else being equal, the larger the sample, the more certainty we have about the effect estimate and the smaller the standard error. When the standard error for an effect size was not provided by the author(s), we calculated it using the formula provided by Hedges and Olkin (1985, p. 86).
4. **Averaging weighted effect sizes.** To determine the average effect on a particular outcome across all relevant studies, we calculated the pooled effect size by estimating a standard random-effects meta-analytic model. The random-effects model tends to produce more conservative effect size estimates (i.e., estimates closer to 0) compared with a fixed-effects model. We estimated the models using a robust variance estimation method (Hedges, Tipton, & Johnson, 2010; Tanner-Smith, Tipton, & Polanin, 2016), which account for both the varying degrees of precision across studies and the nonindependence of effect sizes within studies. The robust variance estimation method upweights effect sizes that are estimated with greater precision and downweights estimates from studies that contribute multiple nonindependent effect size estimates.
5. **Calculating the confidence interval for the average effect.** Even across multiple studies, there remains some uncertainty about the true magnitude of the effect for the population. Another statistic—the 95% confidence interval—gives the upper and lower bounds within which the true population effect is likely to lie. The narrower the interval, the more confident we are about the average effect size accurately reflecting the true effect for the population. The upper and lower bounds of the confidence interval were calculated as the average effect size ± 1.96 times the standard error of the effect size estimate. In this chapter, we use figures like the one below to indicate the magnitude of an average effect and the corresponding 95% confidence interval.



6. **Identifying potential moderating factors.** To determine whether effect sizes are related to art type or outcome domain, we calculated separate average effect sizes and confidence intervals by art type and outcome domain. We used standard meta-analytic procedures to determine whether the variation in effect sizes across art types and outcome domains exceeds what one would expect due to sampling error alone.

We conducted the meta-analysis using the **robumeta** package in R, which is an open-source software environment for statistical computing and graphics.

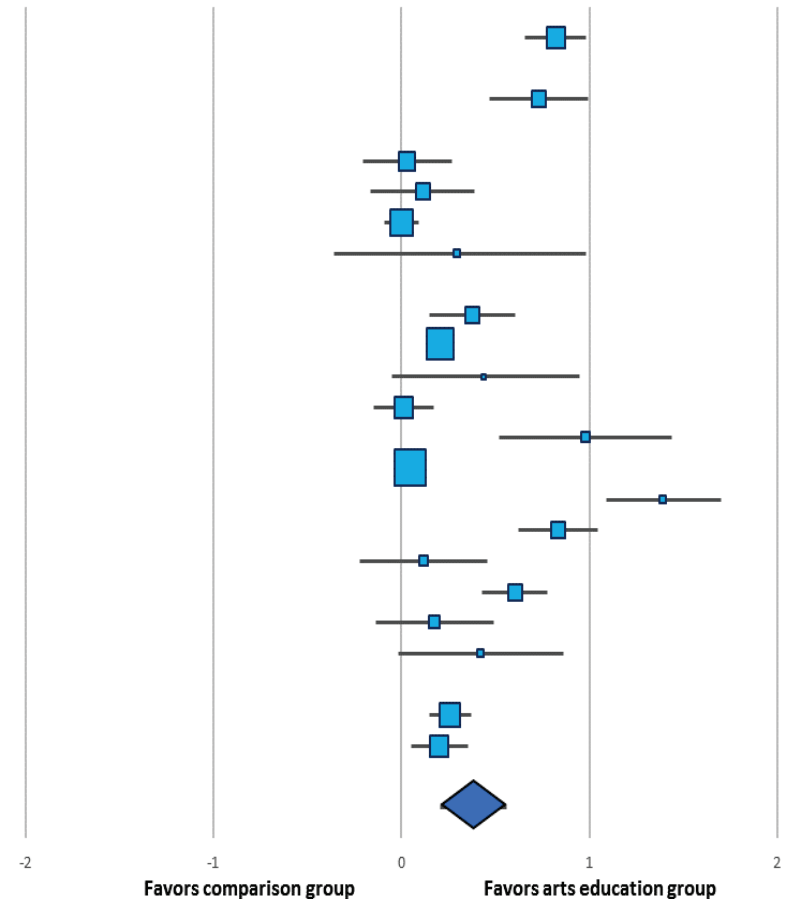
In the sections that follow, we first present the distribution of effect sizes and the average effect size across all the 20 studies included in our meta-analysis, by art type and overall. We then break down the findings according to the type of outcome examined (i.e., academic achievement, art learning, social-emotional learning, and process abilities).

Average Effect of Arts Education Interventions on Student Outcomes

Overall, we found 77 effect size estimates from 20 studies that met the research design requirements for evidence at Tiers I–III. For studies reporting multiple effects, we first computed the average effect size within each study. The average within-study effect sizes from the pool of relevant studies ranged from 0.00 (Kisiel et al., 2006) to 1.40 (Piro & Ortiz, 2009; see Table 4.1). The overall average effect size across all studies was +0.38 (95% confidence interval = 0.20–0.56), a moderate and statistically significant effect (see Table 4.1).

Table 4.1. Average Effects of Arts Education Interventions on Student Outcomes Within and Across Studies, by Art Type and Overall

Art Type and Study Name	Effect Size ^a (Hedges' <i>g</i>)	95% Confidence Interval	
		Lower Limit	Upper Limit
Art (Visual art)			
Catterall & Pepler (2007)	0.82	0.66	0.98
Dance			
Lobo & Winsler (2006)	0.73	0.47	0.99
Drama			
Freeman et al. (2003)	0.03	-0.21	0.27
Goldstein & Winner (2012)	0.11	-0.16	0.39
Kisiel et al. (2006)	0.00	-0.09	0.09
Lan & Morgan (2003)	0.30	-0.38	0.96
Music			
Gromko (2005)	0.38	0.15	0.61
Helmrich (2010)	0.21	0.18	0.24
Henry (2004)	0.44	-0.07	0.93
Hogan et al. (2018)	0.01	-0.15	0.17
Johnson (2011)	0.98	0.52	1.44
Miksza (2010)	0.05	0.02	0.07
Piro & Ortiz (2009)	1.40	1.09	1.70
Rauscher & Zupan (2000)	0.83	0.62	1.05
Register (2004)	0.12	-0.22	0.46
Ritblatt et al. (2013)	0.60	0.43	0.78
Rutkowski & Miller (2003)	0.18	-0.14	0.49
Sharpe (2013)	0.42	-0.02	0.86
Multiple types of art			
Chen et al. (2017)	0.26	0.15	0.37
Horowitz (2016)	0.20	0.05	0.35
Weighted average effect^a	0.38	0.20	0.56



Note. The square for each study represents the average effect size across all relevant outcomes for the study, and the size of the square is proportional to the sample size. The horizontal line for each study shows the 95% confidence interval (i.e., amount of uncertainty) for the average effect. Squares with no lines indicate very precise estimates. Average effects with horizontal lines crossing the vertical line for zero are not statistically significant at the .05 level.

^a The effect size for each study represents a weighted average of all effect sizes reported in the study using the inverse of the variance of each effect estimate as the weight. The overall weighted average effect was estimated using a random-effects model with robust variance estimation.

Source: Authors' analysis of effects from studies meeting design requirements for ESSA Tiers I–III.

Table 4.2. Average Effects of Arts Education Interventions, by Art Type and Type of Student Outcomes

Type of Art and Outcome	Number of Effects	Number of Studies	Average Effect Size (Hedge's g)	Improvement Index	95% Confidence Interval		Favors Comparison Group		Favors Arts Intervention Group									
							-1.0	-0.50	0	+0.50	+1.0	+1.5	+2.0					
Visual art	5	1	0.82	29	0.66	to	0.98											
Social-emotional learning	3	1	0.58	22	0.36	to	0.79											
Process abilities	2	1	1.13	37	0.89	to	1.38											
Dance	6	1	0.73	27	0.47	to	0.99											
Social-emotional learning	6	1	0.73	27	0.47	to	0.99											
Drama	24	4	0.09	4	-0.10	to	0.29											
Art learning	1	1	0.30	12	-0.35	to	0.95											
Social-emotional learning	23	3	0.04	2	-0.09	to	0.18											
Music	38	12	0.45	17	0.17	to	0.72											
Academic achievement	16	6	0.39	15	-0.08	to	0.85											
Art learning	6	3	0.57	21	-0.50	to	1.63											
Social-emotional learning	10	3	0.20	8	-0.62	to	1.03											
Process abilities	6	1	0.83	30	0.62	to	1.05											
Multiple types of art	4	2	0.24	9	-0.04	to	0.52											
Academic achievement	2	1	0.23	9	-0.01	to	0.47											
Art learning	1	1	0.26	10	0.15	to	0.37											
Social-emotional learning	1	1	0.18	7	0.01	to	0.34											
All art types	77	20	0.38	15	0.20	to	0.56											
Academic achievement	18	7	0.44	17	0.00	to	0.88											
Art learning	8	5	0.43	17	0.00	to	0.87											
Social-emotional learning	43	9	0.23	9	-0.01	to	0.47											
Process abilities	8	2	0.93	32	-0.74	to	2.60											

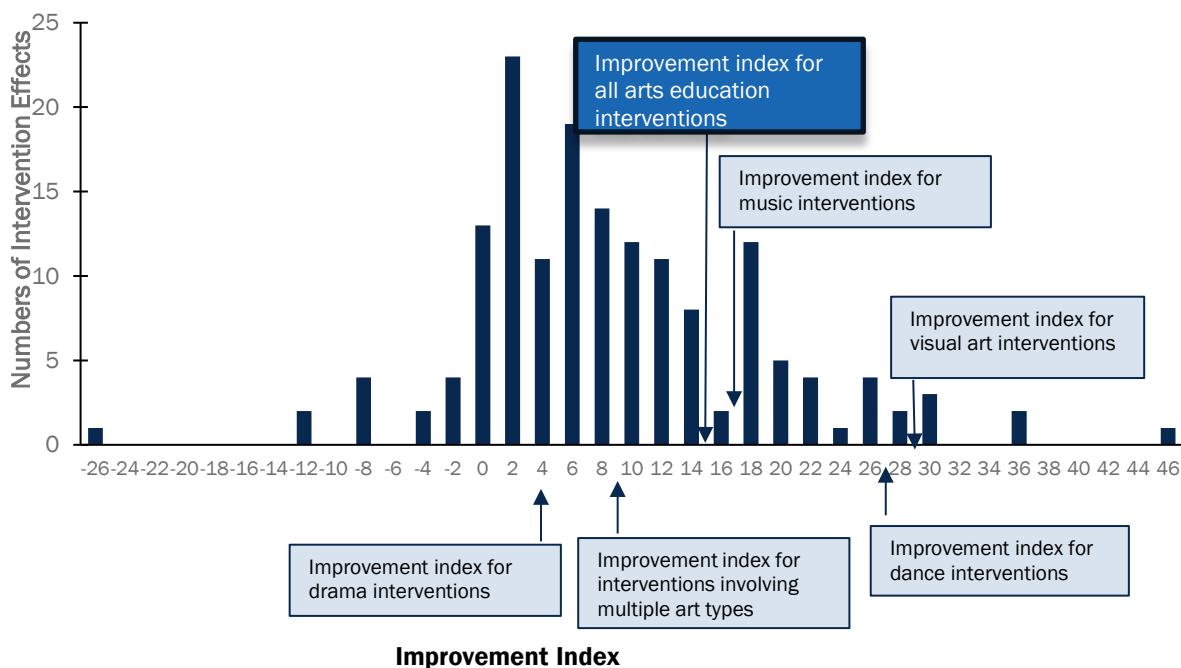
Note. The midpoint of each diamond indicates the point estimate for the average effect; the width of the diamond represents the 95% confidence interval. Diamonds that cross the vertical line for 0 are not statistically significant at the .05 level. When multiple studies provided evidence for an art type and outcome domain combination (i.e., each row in the table), a weighted average effect was calculated based on a random-effects model with robust variance estimation methods. Improvement index indicates the percentile point growth that would be expected for a student at the 50th percentile in the comparison group, had the student received the intervention.

Source: Authors' analysis of effects from studies that meet design requirements for ESSA Tiers I-III.

Average Effects for Visual Art Interventions

Only one study (Catterall & Pepler, 2007) on visual art interventions meets the design requirements for ESSA Tiers I–III. The study produced five effect sizes eligible for inclusion in the meta-analysis: three on social-emotional learning outcomes (self-efficacy) and two on process abilities (originality). The average effect for each outcome domain is positive and statistically significant (Table 4.2). The average effect across all outcomes is also positive and statistically significant with a large magnitude ($g = 0.82$). Another way of looking at the magnitude of this effect is to consider what exposure to the intervention might mean to an average student in the comparison group. With an effect size of 0.82, such a student would likely move from the 50th percentile to the 79th percentile in the outcome (or an improvement index²⁰ of 29 percentile points) had the student received the intervention. Across the 162 intervention effects on relevant outcome domains reported for the 70 interventions in mathematics, reading, and science reviewed by the WWC, the improvement indices ranged from negative 25 percentile points to positive 46 percentile points (see Figure 4.1 and further explanation in Appendix A). The 29-percentile-point increase would put the average effect of the visual art intervention at the 96th percentile among the 70 interventions reviewed by the WWC.

Figure 4.1. Distribution of Improvement Indices for Intervention Effects on Student Achievement in Mathematics, Reading, and Science Based on WWC Intervention Reports



Source: Authors' analyses and improvement indices found in WWC intervention reports.

²⁰ The WWC translates effect sizes into an improvement index, which reflects the percentile point gain that an average student (i.e., the student at the 50th percentile) in the comparison group would experience had he or she been exposed to the intervention.

Average Effect for Dance Interventions

One study of a dance intervention (Lobo & Winsler, 2006) provided six effects eligible for inclusion in the meta-analysis, all in the social-emotional learning outcome domain. This study (Lobo & Winsler, 2006) is an RCT examining the effects of an 8-week instructional program in creative dance/movement on the social competence of low-income preschool children. The study found significantly greater positive improvements over time in children's social competence and both internalizing and externalizing behavior problems for the experimental group compared with the control group ($g = 0.73$ for the average effect; improvement index = 27).

Average Effect for Drama Interventions

Four studies of drama interventions were included in the meta-analysis. One study examined students' art learning outcomes (Lan & Morgan, 2003) and three studies examined students' social-emotional outcomes (Freeman et al., 2003; Goldstein & Winner, 2012; Kisiel et al., 2006). These four studies produced a total of 24 effect sizes for the meta-analysis. The average effect across outcomes was not statistically significant ($g = 0.09$; improvement index = 4). Specifically, the only study that examined the effect of drama education on students' art learning found a positive but nonsignificant effect ($g = 0.30$; Lan & Morgan, 2003). The three studies that examined the effects of drama education on students' social-emotional learning outcomes had a small and nonsignificant average effect ($g = 0.04$; Freeman et al., 2003; Goldstein & Winner, 2012; Kisiel et al., 2006).

Average Effect for Music Interventions

A total of 38 effects from 12 studies of music interventions were included in the meta-analysis. When averaged across outcomes and interventions, the effect of music interventions on student outcomes is moderate and statistically significant ($g = 0.45$; improvement index = 17). One study (Rauscher & Zupan, 2000) examined the effects of a classroom music instruction featuring the keyboard on the process abilities (spatial-temporal reasoning) of kindergarten children. The study found the experimental group scored significantly higher than the comparison group on spatial-temporal tasks after 4 and 8 months of lessons ($g = 0.83$; improvement index = 30). Three studies each examined the effects of music education interventions on students' art learning (Henry, 2004; Johnson, 2011; Rutkowski & Miller, 2003) and social-emotional learning outcomes (Hogan et al., 2018; Miksza, 2010; Ritblatt et al., 2013), and six studies (Gromko, 2005; Helmrich, 2010; Miksza, 2010; Piro & Ortiz., 2009; Register, 2004; Sharpe, 2013) examined student achievement outcomes. When averaged across studies within each outcome domain, all three average effects were positive (ranging from 0.20 to 0.57), but none were statistically significant.

Average Effect for Interventions Involving Multiple Types of Art

Two studies included in the meta-analysis examined interventions that involved multiple types of art. Chen et al.'s (2017) study examined the effect of a professional development program focused on formative assessment practices in arts on students' achievement in multiple art disciplines (music, visual arts, drama, and dance) in 75 schools across the New York city. The study found the intervention had a statistically significant positive effect ($g = 0.26$; improvement index = 10) on students' arts achievement. Horowitz's (2016) study assessed the impact of the Everyday Arts for Special Education program on special education students' academic

achievement and social-emotional behavior in 10 schools in New York's District 75. The study found a statistically significant average effect on social-emotional learning outcomes ($g = 0.18$; improvement index = 7), but the average effect on academic achievement in math and reading was not statistically significant ($g = 0.23$; improvement index = 9). The average effect on student outcomes across the two studies and outcomes also fell short of statistical significance ($g = 0.24$; improvement index = 9).

Summary

The average effect found in the 20 well-designed and well-implemented studies (i.e., studies meeting the design criteria for Tiers I–III) of arts education interventions was statistically significant and moderate in magnitude. Based on the average effect across all the 77 effect sizes and the 20 studies reviewed, one can expect an average child to gain 15 percentile points in a relevant outcome as a result of an arts education intervention ($g = 0.38$). Our analysis also indicated that the effects of arts education interventions varied by art types as well as by outcome domains.²¹ Although these findings suggest that arts education interventions overall may be a promising way to improve student outcomes, readers should exercise caution in interpreting the effects for some art types and outcome domains that are based on a single study or very few studies. Readers also should keep in mind that the studies focus on different types of outcomes for students, which may benefit differently from arts education interventions.

²¹ We conducted omnibus *F*-tests to test whether subgroup means (e.g., means for different art types or means for different outcome domains) are equal.

References

- Abedin, G. (2010). *Exploring the potential of art-based education for adolescents with learning disabilities: A case study of engagement in learning through the arts* (Doctoral dissertation). Retrieved from https://drum.lib.umd.edu/bitstream/handle/1903/10505/Abedin_umd_0117E_11421.pdf
- Abril, C. R. (2006). Learning outcomes of two approaches to multicultural music education. *International Society for Music Education, 24*(1), 30–42.
- Amorino, J. S. (2009). The artistic impetus model: A resource for reawakening artistic expression in adolescents. *Studies in Art Education: A Journal of Issues and Research, 50*(3), 214–231.
- Arts Education Partnership. (2004). *The arts and education: New opportunities for research*. Washington, DC: Author.
- Asbury, C., & B. Rich, B. (Eds.). (2008). *Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition* (Report on Progress in Brain Research). New York, NY: Dana Press.
- Barnes, G. P. (2010). *Moments of meeting: Difficulties and developments in shared attention, interaction, and communication with children with autism during two years of music therapy in a public preschool class* (Doctoral dissertation, Lesley University). Available from ProQuest Dissertations and Theses database. (UMI No. 3449507)
- Benton, M. (2000). *Improving student attitudes and achievement in art* (Master's thesis, St. Xavier University). Retrieved from <https://files.eric.ed.gov/fulltext/ED444921.pdf>
- Bernstein, R. E., Ablow, J. C., Maloney, K. C., & Nigg, J. T. (2014). Piloting PlayWrite: Feasibility and efficacy of a playwriting intervention for at-risk adolescents. *Journal of Creativity in Mental Health, 9*(4), 446–467.
- Bonbright, J., & Faber, R. (2004). *Research priorities for dance education: A report to the nation*. Bethesda, MD: National Dance Education Organization. Retrieved from <https://tinyurl.com/y73zhsyy>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Bowen, R. C. (2010). *Effects of pre-kindergarten music instruction on kindergarten reading and math scores for low SES ELL students* (Doctoral dissertation, Trevecca Nazarene University). Retrieved from <https://pqdtopen.proquest.com/doc/1370975201.html?FMT=AI>
- Bradley, K., Bonbright, J., & Dooling, S. (2013). *Evidence: A report on the impact of dance in the K-12 setting*. Silver Spring, MD: National Dance Education Organization. Retrieved from <https://tinyurl.com/y772ar4z>

- Brock, D., & Lambeth, D. (2013). The effects of music on basic mathematics fact fluency for third grade students. *Cumhuriyet International Journal of Education*, 2(2), 43–60.
- Castro, J. C. (2012). Learning teaching and art through social media. *Studies in Art Education*, 53(2), 152–169.
- Catterall, J. S. (2009). *Doing well and doing good by doing art: The effects of education in the visual and performing arts on the achievements and values of young adults*. Los Angeles, CA: Imagination Group.
- Catterall, J. S., & Peppler, K. A. (2007). Learning in the visual arts and the worldviews of young children. *Cambridge Journal of Education*, 37(4), 543–560.
- Chen, F., Lui, A. M., Andrade, H., Valle, C., & Mir, H. (2017). Criteria-referenced formative assessment in the arts. *Educational Assessment, Evaluation and Accountability*, 29, 297–314.
- Chou, A., & Shih, J. (2010). Show me what you see: An exploration of learning in museums and learning in theatre. *Journal for Learning through the Arts*, 6(1).
- Clements, D. H., & Sarama, J. (2003). Young children and technology: What does the research say? *Young Children*, 58(6), 34–40.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.
- Coulson, A. N., & Burke, B. M. (2013). Creativity in the elementary music classroom: A study of students' music perceptions. *International Journal of Music Education*, 31(4), 428–441.
- Cremata, R., & Powell, B. (2017). Online music collaboration project: Digitally mediated, deterritorialized music education. *International Journal of Music Education*, 35(2), 302–315.
- Darrow, A., Novak, J., Swedberg, O., Horton, M., & Rice, B. (2009). The effect of participation in a music mentorship program on the self-esteem and attitudes of at-risk students. *Australian Journal of Music Education*, 2, 5–16.
- Davis, S. (2016). *Exploring the effects of art-making on racial climate in a multicultural classroom* (Master's thesis, Moore College of Art & Design). Retrieved from <https://files.eric.ed.gov/fulltext/ED567779.pdf>
- Deasy, R. J. (Ed.). (2002). *Critical links: Learning in the arts and student academic and social development*. Washington, DC: Arts Education Partnership. Retrieved from <http://www.eric.ed.gov/PDFS/ED466413.pdf>

- U.S. Department of Education. (2017). Definitions and selection criteria that apply to direct grant programs, § 77.1, 34 C.F.R. pts. 75 and 77. *Federal Register*, 82(145), 35445–35451. Retrieved from <https://www.gpo.gov/fdsys/pkg/FR-2017-07-31/pdf/2017-15989.pdf>
- Elpus, K. (2013). Is it the music or is it selection bias? A nationwide analysis of music and non-music students' SAT scores. *Journal of Research in Music Education*, 61(2), 175–194.
- Ester, D., & Turner K. (2009). The impact of a school loaner-instrument program on the attitudes and achievement of low-income music students. *Contributions to Music Education*, 36(1), 53–71.
- Ferry, L. M. (2016). *Turning anxiety into creativity: Using postmodern principles to alleviate anxiety and stress in the art curriculum and beyond* (Master's thesis, Moore College of Art & Design). Retrieved from <https://files.eric.ed.gov/fulltext/ED567777.pdf>
- Fiske, E. (1999). *Champions of change: The impact of the arts on learning*. Washington, DC: Arts Education Partnership.
- Fleming, M. (2012). *The arts in education, An introduction to aesthetics, theory and pedagogy*. New York, NY: Routledge.
- Freeman, G. D., Sullivan, K., & Fulton, C. R. (2003). Effects of creative drama on self-concept, social skills, and problem behavior. *The Journal of Educational Research*, 96(3), 131–138.
- Frewen, K. G. (2010). Effects of familiarity with a melody prior to instruction on children's piano performance accuracy. *Journal of Research in Music Education*, 57(4), 320–333.
- Gault, B. (2002). Effects of pedagogical approach, presence/absence of text, and developmental music aptitude on the song performance accuracy of kindergarten and first-grade students. *Bulletin of the Council for Research in Music Education*, Spring(152), 54–63.
- Goldstein, T. R., & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, 13(1), 19–37.
- Grant, S., Hamilton, L., Wrabel, S. L., Gomez, C. J., Whitaker, A., Leschitz, J. T., ... Ramos, A. (2017). *Social and emotional learning interventions under the Every Student Succeeds Act: Evidence review*. Washington, DC: RAND.
- Green, D., Mitchell, T., & Taylor, P. (2011). Mentoring in the art classroom. *Improving Schools*, 14(2), 117–129.
- Greene, J. P., Erickson, H. H., Watson, A. R., & Beck, M. I. (2017). The play's the thing: Experimentally examining the social and cognitive effects of school field trips to live theater performances. *Educational Researcher*, 47(4), 246–254.
- Gromko, J. E. (2005). The effect of music instruction on phonemic awareness in beginning readers. *Journal of Research Musical Education*, 53(3), 199–209.

- Gustafson-Hinds, M. A. (2010). *The effectiveness of a unit study—Technology approach* (Doctoral dissertation). Retrieved from <http://irl.umsl.edu/dissertation/486>
- Harding, M., & Haven, B. (2009). The effects of peer coaching in the secondary arts classroom: Intentional watching in the dance studio. *Journal of Dance Education*, 9(2), 41–51.
- Harper, R. (2003). *Enhancing creativity and thinking skills in studio art at secondary level* (Master's thesis, St. Xavier University). Retrieved from <https://files.eric.ed.gov/fulltext/ED479394.pdf>
- Hayes, S. R. (2002). *Improving student performance in fifth grade band through the use of student self-monitoring* (Master's thesis, St. Xavier University). Retrieved from <https://files.eric.ed.gov/fulltext/ED471585.pdf>
- Hedges, L. V. (2009). Effect sizes in nested designs. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 337–355). New York, NY: Russell Sage Foundation.
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65.
- Helmrich, B. H. (2010). Window of opportunity? Adolescence, music, and algebra. *Journal of Adolescent Research*, 25(4), 557–577.
- Hendricks, K. S. (2010). Investing time: Teacher research observing the influence of music history and theory lessons upon student engagement and expressive performance of an advanced high school string quartet. *Bulletin of the Council for Research in Music Education*, Spring(184), 65–78.
- Henry, M. L. (2004). The use of targeted pitch skills for sight-singing instruction in the choral rehearsal. *Journal of Research in Music Education*, 52(3), 206–217.
- Herman, R., Gates, S. M., Arifkhanova, A., Bega, A., Chavez-Herrerias, E. R., Han, E., Harris, M., ... Wrabel, S. (2016). *School leadership interventions under the Every Student Succeeds Act: Evidence review*. Washington, DC: RAND Corporation. Retrieved from <http://www.wallacefoundation.org/knowledge-center/Documents/School-Leadership-Interventions-ESSA-Evidence-Review.pdf>
- Hewitt, M. P. (2001). The effects of modeling, self-evaluation, and self-listening on junior high instrumentalists' music and performance and practice attitude. *Journal of Research in Music Education*, 49(4), 307–322.

- Hewitt, M. P. (2011). The impact of self-evaluation Instruction on student self-evaluation, music performance, and self-evaluation accuracy. *Journal of Research in Music Education*, 59(1), 6–20.
- Hodges, D., & Luehrsen, M. (2010). The impact of a funded research program on music education policy. *Art Education Policy Review*, 111(2), 71–78.
- Hogan, J., Cordes, S., Holochwost, S., Ryu, E., Diamond, A., & Winner, E. (2018). Is more time in general music class associated with stronger extra-musical outcomes in kindergarten? *Early Childhood Research Quarterly*, 45, 238–248.
- Holloway, D. L., & LeCompte, M. D. (2001). Becoming somebody! How art programs support positive identity for middle school girls. *Education and Urban Society*, 33(4), 388–408.
- Horowitz, R. (2016). *District 75, New York City Department of Education Everyday Arts for Special Education impact evaluation*. New York, NY: U.S. Department of Education, Office of Innovation and Improvement.
- Housen, A. (2001, April). *Methods for assessing transfer from an art-viewing program*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.
- Johnson, C. M., & Memmott, J. E. (2006). Examination of relationships between participation in school music programs of differing quality and standardized test results. *Journal of Research in Music Education*, 54(4), 293–307.
- Johnson, D., & Davis, V. W. (2016). The effects of musical ensembles-in-residence on elementary students' auditory discrimination and spatial reasoning: A longitudinal study. *Visions of Research in Music Education*, 28.
- Johnson, D. C. (2011). The effect of critical thinking instruction on verbal descriptions of music. *Journal of Research in Music Education*, 59(3), 257–272.
- Johnson, E. (2017). The effect of symmetrical and asymmetrical peer-assisted learning structures on music achievement and learner engagement in seventh-grade band. *Journal of Research in Music Education*, 65(2) 163–178.
- Kisiel, C., Blaustein, M., Spinazzola, J., Schmidt, C. S., Zucker, M., & Van der Kolk, B. (2006). Evaluation of a theater-based youth violence prevention program for elementary school children. *Journal of School Violence*, 5(2), 19–36.
- Kisida, B., Morrison, B., & Tuttle, L. (2017). *To elevate the role of arts education, measure it*. Washington DC: Brookings. Retrieved from <https://www.brookings.edu/research/to-elevate-the-role-of-arts-education-measure-it/>

- Koops, L. H., Hankins, E. A., Scalise, D., & Schatt, M. D. (2014). Rock orchestra alumni reflections on the impact of participation in The Lakewood Project. *Research Studies in Music Education, 36*(2), 147–164.
- Kraus, N., Hornickel, J., Strait, D. L., Slater, J., & Thompson, E. (2014). Engagement in community music classes sparks neuroplasticity and language development in children from disadvantaged backgrounds. *Frontiers in Psychology, 5*, Article 1403, 1–9.
- Krensky, B. (2001). Going on beyond zebra: A middle school and community-based arts organization collaborate for change. *Education and Urban Society, 33*(4), 427–444.
- Lan, W. Y., & Morgan, J. (2003). Videotaping as a means of self-monitoring to improve theatre students' performance. *The Journal of Experimental Education, 71*(4), 371–381.
- Leijen, A., Admiraal, W., Wildschut, L., & Simons, P. R. (2008). Students' perspectives on e-learning and the use of a virtual learning environment in dance education. *Research in Dance Education, 9*(2), 147–162.
- Lobo, Y. B., & Winsler, A. (2006). The effects of a creative dance and movement program on the social competence of head start preschoolers. *Social Development, 15*(3), 501–519.
- Lopatovska, I., Hatoum, S., Waterstraut, S., Novak, L., & Sheer, S. (2016). Not just a pretty picture: Visual literacy education through art for children. *Journal of Documentation, 72*(6), 1197–1227.
- Lovett, A. (2014). *The relationships among the fine arts, school culture, and high school graduation Rates in Georgia* (Doctoral dissertation, Mercer University). Available from ProQuest Dissertations and Theses database. (UMI No. 3581282)
- Lovin, E., & Lambeth, D. T. (2014). Effects of information and communication technology on engagement and art production for eighth-grade students. *Journal on School Educational Technology, 10*(3), 13–25.
- Ludwig, M., Boyle, A., & Lindsay, J. (2017). *Review of evidence: Arts integration research through the lens of the Every Student Succeeds Act*. New York, NY: The Wallace Foundation.
- Mastorilli, T. M., Harnett, S., & Zhu, J. (2014). Arts Achieve, impacting student success in the arts: Preliminary findings after one year of implementation. *Journal for Learning through the Arts, 10*(1), 1–24.
- Menard, E. A. (2015). Music composition in the high school curriculum: A multiple case study. *Journal of Research in Music Education, 63*(1), 114–136.
- Miksza, P. (2010). Investigating relationships between participation in high school music ensembles and extra-musical outcomes: An analysis of the Education Longitudinal Study of 2002 using bio-ecological development model. *Bulletin of the Council for Research in Music Education, 186*, 7–25.

- Minton, S. (2003). Assessment of high school students' creative thinking skills: A comparison of dance and nondance classes. *Research in Dance Education*, 4(1), 31–49.
- Morrison, S. J., Montemayor, M., & Wiltshire, E. S. (2004). The Effect of a recorded model on band students' performance self-evaluations, achievement and, attitude. *Journal of Research in Music Education*, 52(2), 116–129.
- National Center for Education Statistics. (2017). *The Nation's Report Card: 2016 arts assessment at grade 8*. Washington, DC: Author. Retrieved from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017087>
- National Endowment for the Arts. (1994). *Arts education research agenda for the future*. Washington, DC: Author.
- Neddeau, B. M. (2013). *The relations among school status variables, teacher academic and arts curricular emphases, and student academic achievement in grades 1, 3, and 5* (Doctoral dissertation). Retrieved from <https://repository.usfca.edu/diss/91/>
- Neel, M. M. (2017). Mariachi and Spanish speaking English learners: District initiatives, models, and education policy. *Arts Education Policy Review*, 118(4), 208–219.
- Neville, H. J., Andersson, A., Bagdade, O., Bell, T., Currin, J., Fanning, J., ... Yamada, Y. (2008). Effects of music training on brain and cognitive development in under-privileged 3- to 5-year-old children: Preliminary results. In C. Asbury, & B. Rich (Eds.), *Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition* (pp. 105–116). (Report on Progress in Brain Research). New York, NY: Dana Press.
- Paney, A. S., & Kay, A. C. (2015). Developing singing in third-grade music classrooms: The effect of a concurrent-feedback computer game on pitch-matching skills. *National Association for Music Education*, 34(1), 42–49.
- Parisi, J. (2004). Fourth- and fifth-grade students' affective response and ability to discriminate between melody and improvisation after receiving instruction in singing and/or playing a piece in the blues style. *International Journal of Music Education*, 22(1), 77–86.
- Peppler, K. (2013). *New opportunities for interest-driven arts learning in a digital age*. New York, NY: The Wallace Foundation. Retrieved from <https://www.wallacefoundation.org/knowledge-center/Documents/New-Opportunities-for-Interest-Driven-Arts-Learning-in-a-Digital-Age.pdf>
- Piro, J. (2009, June). Music training and literacy development. *Literacy Today*, 32–34.
- Piro, J. M., & Ortiz, C. (2009). The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. *Psychology of Music*, 37(3), 325–347.
- Rapp-Paglicci, L., Stewart, C., & Rowe, W. (2012). Improving outcomes for at-risk youth: Findings from the Prodigy Cultural Arts Program. *Journal of Evidence-Based Social Work*, 9(5), 512–523.

- Rauscher, F. H., & Zupan, M. A. (2000). Classroom keyboard instruction improves kindergarten children's spatial-temporal performance: A field experiment. *Early Childhood Research Quarterly, 15*(2), 215–228.
- Rawlings, J. R., & Stoddard, S. A. (2017). Peer connectedness in the middle school band program. *Research Studies in Music Education, 39*(1), 121–135.
- Register, D. (2004). The effects of live music groups versus an educational children's television program on the emergent literacy of young children. *Journal of Music Therapy, 41*(1), 2–27.
- Reifinger, J. L. (2012). The acquisition of sight-singing skills in second-grade general music: Effects of using solfège and of relating tonal patterns to songs. *Journal of Research in Music Education, 60*(1), 26–42.
- Riley, P. E. (2006). Including composition in middle school band: Effects on achievement, performance, and attitude. *Update: Applications of Research in Music Education, 25*(1), 28–38.
- Ritblatt, S., Longstreth, S., Hokoda, A., Cannon, B., & Weston, J. (2013). Can music enhance school-readiness socioemotional skills? *Journal of Research in Childhood Education, 27*, 257–266.
- Rutkowski, J., & Miller, M. S. (2003). The effect of teacher feedback and modeling on first graders' use of singing voice and developmental music aptitude. *Bulletin of the Council for Research in Music Education, 156*, 1–10.
- Seidel, S., Tishman, S., Winner, E., Hetland, L., & Palmer, P. (2009). *The qualities of quality: Understanding excellence in arts education*. Cambridge, MA: Project Zero. Retrieved from <https://www.wallacefoundation.org/knowledge-center/Documents/Understanding-Excellence-in-Arts-Education.pdf>
- Sharpe, N. N. (2013). *The relationship between music instruction and academic achievement in mathematics* (Doctoral dissertation, Walden University). Available from ProQuest Dissertations and Theses database. (UMI No. 3604333)
- Shields, C. (2001). Music education and mentoring as intervention for at-risk urban adolescents: Their self-perceptions, opinions, and attitudes. *Journal of Research in Music Education, 49*(3), 273–286.
- Snow, C. S., & McLaughlin, T. F. (2005). Effects of teaching perspective in a structured and systematic way on still life drawing of elementary students: An empirical study. *Educational Research Quarterly, 28*(3), 17–26.
- Sosin, A. A., Bekkala, E., & Pepper-Sanello, M. (2010). Visual arts as a lever for social justice education: Labor studies in the high school art curriculum. *Journal for Learning through the Arts, 6*(1). Retrieved from <https://files.eric.ed.gov/fulltext/EJ1093573.pdf>

- Southgate, D. E., & Roscigno, V. J. (2009). The impact of music on childhood and adolescent achievement. *Social Science Quarterly*, 90(1), 4–21.
- St. John, P. A. (2006). Finding and making meaning: Young children as musical collaborators. *Psychology of Music*, 34(2), 238–261.
- Stankiewicz, M. A., Amburgy, P. M., & Bolin, P. E. (2004). Questioning the past: Contexts, functions, and stakeholders in 19th-century art education. In E. W. Eisner & M. D. Day (Eds.), *Handbook of research and policy in art education*. London, UK: Routledge. Retrieved from <https://www.routledgehandbooks.com/doi/10.4324/9781410609939.ch2>
- Szekely, G., & Buckman, J. A. (2012). *Art teaching: Elementary through middle school*. New York, NY: Routledge.
- Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex metanalytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology*, 2(1), 85–112.
- Taylor, P. G. (2000). Madonna and hypertext: Liberatory learning in art education. *Studies in Art Education*, 41(4), 376–389.
- Thomas, M. K. (2016). *Music education and its causal impact on student engagement and success: A program evaluation of Little Kids Rock* (Working Paper). Washington, DC: National Endowment for the Arts.
- Thornton, L. (2008). The effect of grade, experience, and listening condition on the melodic error detection of fifth- and sixth-grade woodwind students. *Update: Applications of Research in Music Education*, 26(2), 4–10.
- Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in Psychology*, 4, 1–7.
- U.S. Department of Education. (2016). *Non-regulatory guidance: Using evidence to strengthen education investments*. Washington, DC: Author. Retrieved from <https://www2.ed.gov/policy/elsec/leg/essa/guidanceuseseseinvestment.pdf>
- Varian, S. (2016). *Choosing creatively: Choice-based art education in an inclusive classroom* (Master's thesis, Moore College of Art & Design). Retrieved from <https://files.eric.ed.gov/fulltext/ED567778.pdf>
- Vega, L. (2001). *Increasing student music achievement through the use of motivational strategies* (Master's thesis, St. Xavier University).
- Walker, M. A. (2014). From theory to practice: Concept-based inquiry in a high school art classroom. *Studies in Art Education*, 55(4), 287–299.

- Werner, P. D. (2011). *Effects of synchronous lessons on students in an asynchronous online art course* (Doctoral dissertation, University of Florida). (UMI No. 3514911)
- What Works Clearinghouse (WWC). (2017). *What Works Clearinghouse procedures handbook* (Vol. 4). Washington, DC: Author. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_handbook_v4.pdf
- Winner, E., & Hetland, L. (2000). The arts in education: Evaluating the evidence for a causal link. *The Journal of Aesthetic Education*, 34(3/4), 3–10.
- Winsler, A., Ducenne, L., & Koury, A. (2011). Singing one's way to self-regulation: The role of early music and movement curricula and private speech. *Early Education and Development*, 22(2), 274–304.
- Wolf, D. P., & Holochwost, S. (2014). *Our voices count: The potential impact of strength-based music programs in juvenile justice settings*. Washington, DC: National Endowment for the Arts. Retrieved from http://wolfbrown.com/images/books_reports/documents/ourvoicescount.pdf

Appendix A. Technical Approach to the Evidence Review

This evidence review was conducted in two stages as described in detail below.

Stage 1: Project and Subject Articulation Phase

After consulting members of The Wallace Foundation, seminal works on arts education, and the U.S. Department of Education's guidance for interpreting ESSA's tiers of evidence, we informed our logic model and drew clearer boundaries for this evidence review.

First, we arrived at definitions for arts education and the populations and time periods of interest for this evidence review.

- ▶ We began with an understanding that we were interested in classes or courses in the arts taught in school. The main purpose of such classes would be to learn about and practice a particular type of art. With the emphasis on learning, we defined arts education as follows:
 - Arts education refers to arts lessons or classes offered in prekindergarten through Grade 12 that are (a) standards based and (b) taught by certified arts specialist teachers or teaching artists through (c) an explicit or implied sequential arts curriculum in the (d) subjects of art/visual arts, media arts, music, dance, and drama/theater.

This definition excludes from our review studies of arts integration, which was the focus of a previous evidence review (Ludwig, Boyle, & Lindsay, 2017). We also decided to exclude summaries about research, secondary sources about research, and studies of brain research or therapeutic interventions that were not salient for teaching and learning in school. However, we did include studies about arts teaching and changes in instruction that were expected to affect student outcomes. Many studies investigate the conditions for arts teaching or teaching methods, and ESSA includes opportunities to support professional development for arts teachers.

- ▶ Then, we defined the populations of interest for this review: students in prekindergarten through Grade 12 who are learning academic content in schools located in the United States or its territories.
 - This definition excludes studies conducted with similarly aged students in other countries (unless a U.S. student population was participating in the study); studies that examined interventions or programs conducted outside the school setting; and studies involving students in postsecondary settings.
- ▶ Third, we identified the time period of interest: Study reports eligible for this evidence review must have been published during or after 2000.

As we did for the arts integration evidence review, our review team consulted with partners at The Wallace Foundation and with education policy experts to determine whether to define the tiers of evidence based on the ambiguous parameters specified in the ESSA law itself or to define the tiers based on the additional Non-Regulatory Guidance issued by the U.S. Department of Education (2016). The evidence tiers as defined in the latter document are more stringent in

that the guidance specifies additional requirements for Tier I and Tier II evidence. The stakes involved in adopting an arts education program are high in terms of district and state resources and student learning. For this reason, we chose the more rigorous interpretation of tiers of evidence specified in U.S. Department of Education's (2016) supplemental guidance document.

Stage 2: Searching, Prescreening, Screening, Reviewing, Extracting Study Details, and Synthesizing

Stage 2 involved a number of distinct activities: (a) searching for relevant reports; (b) prescreening abstracts by experienced content screeners; (c) screening in-scope report abstracts and full texts against search criteria; (d) reviewing and classifying studies based on ESSA's tiers of evidence; (e) extracting from each report information about study samples, settings, program components, outcomes, and magnitude of effects; and (f) synthesizing findings by tier and by art type and by meta-analyzing all effects from studies of interventions that met research design requirements for Tiers I–III. These activities are described in more detail below.

Searching for Relevant Reports

We implemented a multipronged search strategy to uncover as many reports on arts education interventions as possible. First, we developed an extensive series of search strings based on a review of relevant reports, professional organization websites, and the National Core Arts Standards. We also scanned the websites of organizations that are involved in arts education policy and research to uncover additional published news articles, policy statements or reports that might guide our literature search. Although we made every effort to systematically locate relevant published reports, we were not able to search developers' websites or reanalyze the studies in published meta-analyses.

Then, working with AIR's research librarian, we ran search strings and delimiters through education-focused databases. We conducted two rounds of search with the understanding that report titles and abstracts were likely to use many combinations of words indicating original research studies while omitting words describing an art type. (See Box A.1 for a list of the research databases searched and the search strings used.) Third, we examined the reference lists in the studies deemed eligible by screeners to identify published studies that met our parameters for population and time period and were missed in our initial database searches. The database searches yielded 15,940 results. After the removal of duplicates, we were left with 7,170 unique results, which grew to 7,405 with the addition of reports found through examining reference lists and multiple sources which were hand-searched.

Many reports that we considered adding to the review database, upon examination, did not meet the search parameters. Some of these reports, however, were good sources of background and history about the development of specific types of the art, the current focus on research and practice in arts education, and examples of teacher practice in arts classes.

Box A.1. Extent of Literature Search

The basic sketch of the arts education literature search:

The Arts terms

AND

Outcomes terms

AND

PreK-12 terms

AND

Study type/terms associated with empirical studies

Search in EBSCO's ERIC, Education Source, and Education Research Complete, on 12/06/2017:

(TI Art OR TI arts OR "arts education" OR "art education" OR TI "visual arts" OR TI "visual art" OR TI painting OR TI photography OR TI music* OR TI chorus OR TI orchestra* OR TI band OR TI bands OR TI theat* OR TI drama OR TI opera OR TI operas OR TI dance* OR TI dancing OR TI choreograph* OR TI ballet* OR AB "graphic arts" OR AB "graphic art" OR AB "media arts" OR AB "media art" OR AB "digital arts" OR AB "digital art" OR "performing art*" OR AB "fine art*" OR TI choir OR "arts instruction" OR "art instruction" OR "creative movement" OR "dance movement" OR "arts learning" OR "artistic learning") in Select a field, optional

AND

("educational opportunit*" OR TI outcome* OR TI Proficien* OR TI skills OR TI motivat* OR TI engagement OR TI Performance OR achievement* OR "academic gains" OR "academic performance" OR "academic behavior" OR "academic behaviors" OR attendance OR TI graduation OR "cognitive outcome*" OR TI cognition OR TI cognitive OR "behavioral outcome*" OR TI behavior* OR "social emotional" OR "social and emotional" OR "school readiness" OR "social competence" OR "social competency" OR "school culture" OR "school climate" OR TI creativity OR AB creativity OR "critical reasoning" OR "student retention" OR "drop out" OR "drop outs" OR TI memory OR TI "analytic" OR TI analytical OR "language proficiency" OR "language development" OR "self-efficacy" OR AB "creative thinking" OR TI culture OR TI cultural OR AB "multi-cultural" OR AB "problem-solving" OR AB "self-concept" OR TI spatial* OR AB spatial* OR Wellbeing OR "Well being" OR TI learning OR "academic development" OR "academic effect*" OR "social development" OR "social effect*" OR TI Reading OR TI writing OR TI literacy OR TI language OR TI math* OR TI numeracy OR AB empathy OR AB tolerance OR AB "self-confidence" OR AB responsibility OR AB "self-management" OR AB resilient* OR AB perseverance OR AB "self-regulation" OR AB "conflict resolution" OR AB "critical thinking" OR AB cooperation OR AB cooperative OR AB comprehension OR AB reflective OR AB reflection OR AB awareness OR AB perception OR AB intelligence OR AB "test scores" OR AB interpersonal OR "arts performance" OR "art performance" OR "artistic performance" OR "arts knowledge" OR "art knowledge" OR "art skills" OR "arts skills" OR TI attitude* OR "attitudes towards art" OR "attitude towards art" OR "attitude toward art" OR "attitudes toward art" OR "art attitudes" OR "arts attitudes" OR bullying OR TI identity OR "Fine motor skill*" OR "non-verbal reasoning" OR "emotional development" OR "concentration" OR "leadership skills" OR "community service" OR "social engagement" OR "expressive ability" OR "expressive abilities" OR "grade attainment" OR "course grade*" OR "grade received" OR "grades received" OR "grade attained" OR AB "grade point average*" OR AB imagination OR TI ability OR TI abilities OR "cultural awareness" OR TI understanding OR TI collaboration OR TI communication OR "civic engagement" OR "cultural understanding") in Select a field, optional

AND

(preschool* OR "pre-school*" OR "pre-K*" OR preK* OR "PK-12*" OR AB "PK-8*" OR AB "PK-6*" OR AB "PK-5*" OR "K-12*" OR AB "K-8*" OR AB "K-6*" OR "K-5*" OR prekindergarten* OR "pre-kindergarten*" OR kindergarten* OR TI students OR TI youth OR AB youth OR TI young OR children OR "high school*" OR AB teen* OR adolescen* OR toddler* OR "secondary school*" OR "junior high" OR "middle school*" OR grade OR AB graders OR TI elementary OR "elementary school*" OR "early childhood" OR "school age*" OR "school based" OR "school system*" OR TI school* OR "school district*" OR TI course* OR TI class* OR extracurricular OR "extra-curricular" OR AB elective* OR AB "young people" OR "after-school" OR afterschool OR TI training OR TI instruction) in Select a field, optional

AND

(AB "treatment" OR AB "comparison" OR AB "comparative" OR AB "control" OR AB "effect size*" OR AB "statistical" OR AB "statistically" OR AB "random" OR AB "randomly" OR TI "randomized" OR AB "randomized" OR AB "sample" OR AB "study" OR AB "experiment" OR AB "experimental" OR AB "quasi-experimental" OR AB "empirical" OR AB correlat* OR TI "meta-analy*" OR AB "meta-analy*" OR AB "relationship between" OR AB "relationships between" OR AB "relationship among" OR AB "relationships among" OR AB measures OR AB measure OR AB research OR AB impact* OR AB benefit OR AB benefits OR AB result OR AB results OR AB resulted OR TI evaluation* OR AB evaluation* OR AB effect* OR AB influence* OR AB significant* OR AB findings OR AB data) in Select a field, optional

Limits: English language; 2000-2018

Additional search in EBSCO's ERIC, Education Source, and Education Research Complete, on 01/04/2018:

(AB "computer aided design*" OR "fashion design*" OR "graphic design*" OR TI drawing* OR AB sculpture* AB sculpting OR AB "studio art*" OR "Movement vocabulary" OR "dance movements" OR "Dance phrase*" OR "Dance sequence*" OR "System of dance" OR "Systems of dance" OR "Musical idea*" OR "Musical understanding" OR "Musical creativity" OR TI Rhythmic OR AB melodic OR AB Melody OR AB Melodies OR "musical tempo*" OR "music tempo*" OR "musical phras*" OR "music phras*" OR "musical interpretati*" OR "musical harmon*" OR "Guided drama*" OR "Acting technique*" OR Playwright* OR Dramaturg* OR "Improvised drama*" OR "Scripted drama*" OR "art exhibit*" OR "art museum*" OR "artistic talent" OR TI "Aesthetics") in Select a field, optional

AND

("educational opportuniti*" OR TI outcome* OR TI Proficien* OR TI skills OR TI motivat* OR TI engagement OR TI Performance OR achievement* OR "academic gains" OR "academic performance" OR "academic behavior" OR "academic behaviors" OR attendance OR TI graduation OR "cognitive outcome*" OR TI cognition OR TI cognitive OR "behavioral outcome*" OR TI behavior* OR "social emotional" OR "social and emotional" OR "school readiness" OR "social competence" OR "social competency" OR "school culture" OR "school climate" OR TI creativity OR AB creativity OR "critical reasoning" OR "student retention" OR "drop out" OR "drop outs" OR TI memory OR TI "analytic" OR TI analytical OR "language proficiency" OR "language development" OR "self-efficacy" OR AB "creative thinking" OR TI culture OR TI cultural OR AB "multi-cultural" OR AB "problem-solving" OR AB "self-concept" OR TI spatial* OR AB spatial* OR Wellbeing OR "Well being" OR TI learning OR "academic development" OR "academic effect*" OR "social development" OR "social effect*" OR TI Reading OR TI writing OR TI literacy OR TI language OR TI math* OR TI numeracy OR AB empathy OR AB tolerance OR AB "self-confidence" OR AB responsibility OR AB "self-management" OR AB resilien* OR AB perseverance OR AB "self-regulation" OR AB "conflict resolution" OR AB "critical thinking" OR AB cooperation OR AB cooperative OR AB comprehension OR AB reflective OR AB reflection OR AB awareness OR AB perception OR AB intelligence OR AB "test scores" OR AB interpersonal OR "arts performance" OR "art performance" OR "artistic performance" OR "arts knowledge" OR "art knowledge" OR "art skills" OR "arts skills" OR TI attitude* OR "attitudes towards art" OR "attitude towards art" OR "attitude toward art" OR "attitudes toward art" OR "art attitudes" OR "arts attitudes" OR bullying OR TI identity OR "Fine motor skill*" OR "non-verbal reasoning" OR "emotional development" OR "concentration" OR "leadership skills" OR "community service" OR "social engagement" OR "expressive ability" OR "expressive abilities" OR "grade attainment" OR "course grade*" OR "grade received" OR "grades received" OR "grade attained" OR AB "grade point average*" OR AB imagination OR TI ability OR TI abilities OR "cultural awareness" OR TI understanding OR TI collaboration OR TI communication OR "civic engagement" OR "cultural understanding") in Select a field, optional

AND

(preschool* OR "pre-school*" OR "pre-K*" OR preK* OR "PK-12*" OR AB "PK-8*" OR AB "PK-6*" OR AB "PK-5*" OR "K-12*" OR AB "K-8*" OR AB "K-6*" OR "K-5*" OR prekindergarten* OR "pre-kindergarten*" OR kindergarten* OR TI students OR TI youth OR AB youth OR TI young OR children OR "high school*" OR AB teen* OR adolescen* OR toddler* OR "secondary school*" OR "junior high" OR "middle school*" OR grade OR AB graders OR TI elementary OR "elementary school*" OR "early childhood" OR "school age*" OR "school based" OR "school system*" OR TI school* OR "school district*" OR TI course* OR TI class* OR extracurricular OR "extra-curricular" OR AB elective* OR AB "young people" OR "after-school" OR afterschool OR TI training OR TI instruction) in Select a field, optional

AND

(AB "treatment" OR AB "comparison" OR AB "comparative" OR AB "control" OR AB "effect size*" OR AB "statistical" OR AB "statistically" OR AB "random" OR AB "randomly" OR TI "randomized" OR AB "randomized" OR AB "sample" OR AB "study" OR AB "experiment" OR AB "experimental" OR AB "quasi-experimental" OR AB "empirical" OR AB correlat* OR TI "meta-analy*" OR AB "meta-analy*" OR AB "relationship between" OR AB "relationships between" OR AB "relationship among" OR AB "relationships among" OR AB measures OR AB measure OR AB research OR AB impact* OR AB benefit OR AB benefits OR AB result OR AB results OR AB resulted OR TI evaluation* OR AB evaluation* OR AB effect* OR AB influence* OR AB significant* OR AB findings OR AB data) in Select a field, optional

Limits: English language; 2000-2018

Prescreening of Report Abstracts for Relevance

For this evidence review, two experienced prescreeners conducted a preliminary screening of the abstract of each of the 7,405 reports resulting from the multilevel search strategy implemented (see Box A.1). The bulk of these reports came as a result of the initial systematic search of relevant education databases. As additional reports were identified through review of bibliographies and websites, the reports were noted, uploaded by the research librarian, and prescreened.

The prescreeners identified reports that were out of scope for the following reasons: (a) the publication date was out of the review timeframe, (b) the study sample did not include students in the United States, (c) the topic was out of scope, or (d) the topic was not relevant to arts education.

Of the 7,405 report abstracts uploaded to our review database, prescreeners identified 1,863 to keep and 5,542 to exclude.

Screening Eligible Reports: Abstracts and Full Texts

To prepare the screeners who would examine the abstracts and full-text reports that passed prescreening, two sessions were organized by the arts education review research and database management team. In these sessions, screeners were trained on how to identify the topic in reports and to differentiate arts education studies from arts integration studies. The training team devoted time to this distinction because the titles and abstracts of reports about arts integration often sound similar to ones about studying the arts in self-contained classes. The training also prepared screeners regarding the use of the term discipline-based arts instruction (DBAI), a national initiative sponsored by the Getty Center for Education in the Arts. Based on readings about trends in arts education, we expected to find reports about this particular initiative or generally referring to arts education as discipline-based arts education. Screeners were also trained on the use of various database functions. An adjudication step was added for the screening phrase. Senior project staff with content expertise in the areas of arts integration and arts education literature were given the responsibility of adjudicating assignments when screeners were not able to decide or disagreed as to whether an article was in scope.

The screening began with the abstracts of reports that passed prescreening. One third of these report abstracts were each assigned to two screeners for two independent rounds of screening. Abstracts that seemed potentially relevant based on the screening were forwarded to the research librarian so that she could obtain the full text of the reports.

Abstract Screening. Trained screeners reviewed the abstracts of the reports that passed the prescreening and answered four questions for each:

1. Does the abstract mention discipline-based arts instruction (DBAI) (or synonyms, such as arts education or arts learning or sequential arts) or does it mention specific instructional approaches to teaching and learning for the arts?
2. Does the abstract report on a study that collected data and reported findings (for example, data from a survey, assessments, interviews, observations)?

3. Does the abstract indicate that the report is a literature review, cites DBAI (or synonyms) studies, or includes a meta-analysis developed by examining the outcomes of other studies?
4. Does the abstract mention students in prekindergarten through Grade 12 who attended schools in the United States or its territories?

If the screeners were not sure of answers to the screening questions, then the study progressed to adjudication by an experienced content screener and eventually to the full-text screening phase, if determined as eligible. Of the reports that passed prescreening and abstract screening, 704 were identified as prospects for full-text review. With additional review, a total of 665 full-text articles were requested from the librarian based on the abstracts.

Full-Text Screening. The research librarian attempted to obtain the full text of all 665 reports that passed the abstract screening and retrieved the full text for all but 40 reports. Thirty-eight of the reports were considered out of scope by experienced screeners; one report was identified as a resource, and one report was not available. The 625 full-text reports were each reviewed by one screener, who was prompted to answer the following questions:

1. Does the report mention DBAI (or synonyms, such as arts education or arts learning or sequential arts) or does it mention specific instructional approaches to teaching and learning for the arts?
2. Does the report describe data that were collected and analyzed about DBAI or synonyms (for example, data from a survey, assessments, interviews, observations)?
3. Does the report describe a sample of students in prekindergarten through Grade 12 who were attending a school in the United States or its territories?
4. Does the report describe student-related outcomes (for example, artistic, academic, behavioral, social-emotional outcomes)?
5. Does the report cite other studies (in a literature review or through a meta-analysis) on DBAI (or synonyms, such as arts education or arts learning or sequential arts) or does it mention specific instructional approaches to teaching and learning for the arts?

For each question, the full-text report screeners answered “yes,” “no,” or “unable to determine.” If the screeners were unsure about the answers to each of these five questions, then the report was forwarded to experienced reviewers for adjudication. Of the 625 full-text reports screened, 339 reports were removed from the pool of potentially eligible arts education reports, leaving 286 reports that were forwarded to WWC-certified reviewers for review based on ESSA’s evidence tiers.

Reviewing and Classifying Studies by WWC-Certified Reviewers

Each relevant full-text report that passed screening was then reviewed by researchers who are certified to review studies using WWC standards. These reviewers classified the studies in the reports into evidence tiers based on U.S. Department of Education’s (2016) Non-Regulatory Guidance document, which recommends that evidence from each study be classified based on (a) the study’s research design, (b) whether the study reports statistically significant findings, (c) the sample size of the study, and (d) the number of sites in which the study was conducted.

The study reviewers made decisions as to whether study evidence fit into Tiers I–III based on those study features.

The classification of study evidence into Tier IV was informed by the principles from the U.S. Department of Education’s (2016) Non-Regulatory Guidance on evidence-based interventions as well as the approach to “unpacking Tier IV” used in the school leadership evidence review conducted by Herman et al. (2016). To determine whether a study of an intervention provided Tier IV evidence, we first examined whether the report(s) associated with that study included a logic model (i.e., a graphic representation of how the intervention was intended to affect relevant outcomes). If a study’s report(s) did not include a graphic logic model, then we determined whether the study’s report(s) described at least a theory of action explaining how at least one feature of the intervention was intended to affect relevant outcomes and whether at least one of the described features was supported by empirical research. We based our decisions about whether a logic model or theory of action was supported by empirical research on (a) the study’s explicit reference to empirical research or (b) our own understanding of relevant empirical research. Those that did not provide any of those types of information were identified as providing “no ESSA tier-aligned evidence.” The criteria for classification of evidence into ESSA tiers are summarized in Table A.1. The final disposition of the 7,405 reports that were originally identified as potentially eligible for the review is shown in Figure A.1.

Table A.1. Criteria for Determining Tier of Evidence Based on ESSA and U.S. Department of Education’s (2016) Non-Regulatory Guidance

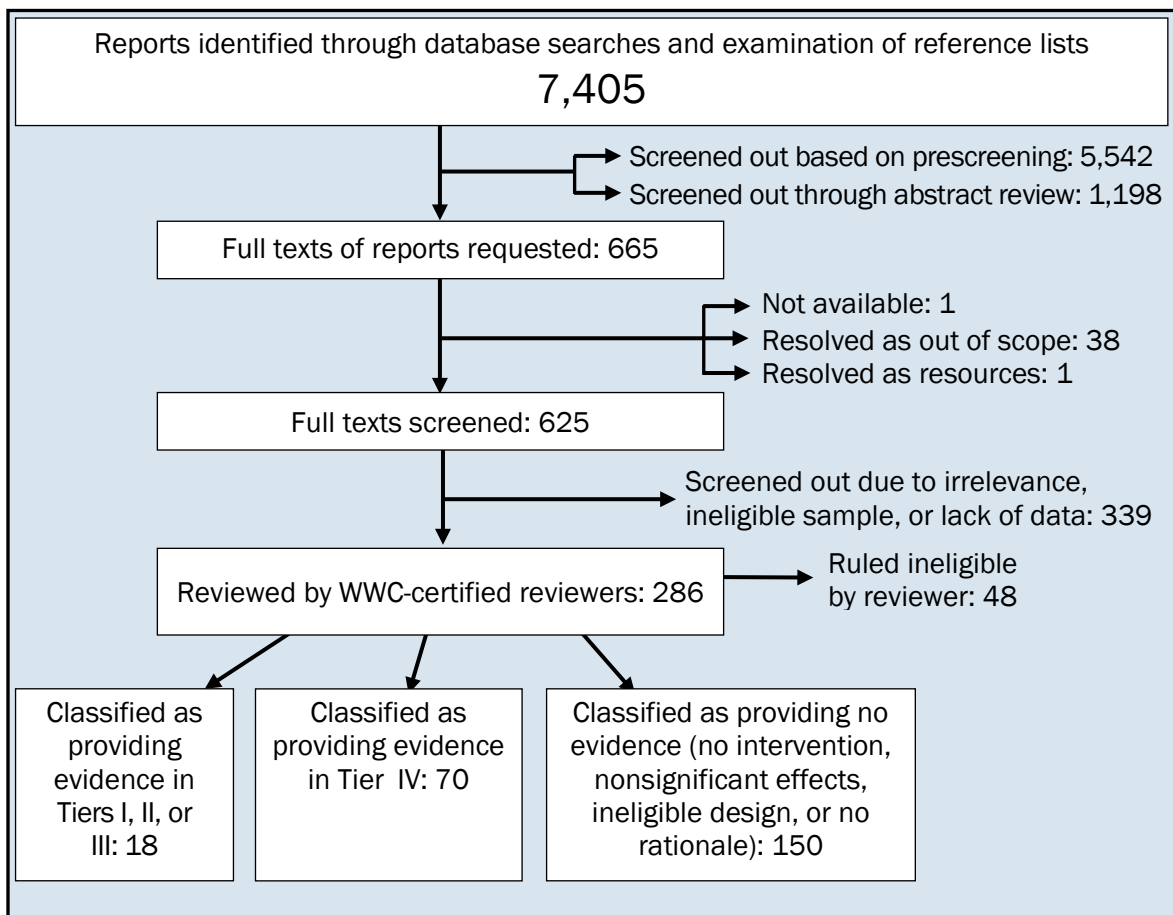
Evidence Tier	Randomized Study	Low Attrition	Quasi-experimental Study	Baseline Equivalence	Correlational Study With Control for Confounds	Sample >349 Students	Multisite Sample	Statistically Significant Findings ^a	Logic Model or Rationale
Tier I	✓	✓				✓	✓	✓	
Tier II	✓		✓	✓		✓	✓	✓	
Tier III	✓		✓		✓			✓	
Tier IV									✓

Note. ESSA = Every Student Succeeds Act. To be eligible for Tiers I–III, studies using group designs must avoid the $N = 1$ confound, such as when a single teacher is responsible for the instruction of students in one condition. In such cases, differences between the conditions can also be attributed to that other factor (in this case, the single teacher) rather than the intervention.

^a For Tiers I–III, in accordance with WWC standards, impact estimates from clustered group design studies—those in which units of assignment were entire clusters rather than individuals—must account for the clustering. If the study-reported findings ignored clustering, WWC would apply a clustering correction to the study reported findings by using an intraclass correlation of .20 for academic achievement outcomes and .10 for attitude outcomes. These corrections were applied in this review as well.

Source: Authors’ review protocol.

Figure A.1. Disposition of Reports of Studies of Arts Education Interventions Found in This Evidence Review



Extracting Study Details

For all arts education interventions (including those for which Tier IV evidence exists), we gathered as much information about the interventions as possible to “potentially increas[e] opportunities for states and districts to develop interventions for their own context or to build evidence on popular but under-researched interventions” (Herman et al., 2016, p. 12). Specifically, the WWC-certified reviewers coded and recorded the characteristics of the report, study (i.e., research design and tier), sample and setting, components of interventions, outcomes, and effect sizes (see Table A.2 for a list of these features). The codes of study features were used in the synthesis of evidence across studies.

Table A.2. Types of Information Recorded From Each Eligible Study Report

Report	Study	Sample and Setting	Intervention	Effect Size
<ul style="list-style-type: none"> ▶ Year of publication ▶ Publication vehicle 	<ul style="list-style-type: none"> ▶ Whether study meets WWC standards ▶ ESSA evidence tier 	<ul style="list-style-type: none"> ▶ Sample sizes ▶ Grade levels ▶ Race/ethnicity characteristics ▶ Students' socioeconomic status ▶ Setting 	<ul style="list-style-type: none"> ▶ Single or multiple components ▶ Teacher professional development ▶ Involvement of certified art teachers and professional artists ▶ Types of program materials used ▶ Student field trip ▶ Schoolwide model 	<ul style="list-style-type: none"> ▶ Outcome of interest ▶ Effect size (<i>g</i>)

Synthesizing of Study Evidence

Findings from studies providing evidence in ESSA Tiers I–IV were synthesized in two ways. First, we organized the ESSA-related findings using tables cross-tabulating studies providing different tiers of evidence by type of art and outcome domain. Second, we meta-analyzed effect sizes to arrive at an average effect across all 20 studies that met the research design requirements for evidence at Tiers I–III and average effects for studies examining different types of art and outcomes. Below, we provide further details about the meta-analysis we conducted for this evidence review.

Meta-Analytic Methods

A meta-analysis is a set of statistical procedures used to combine the findings from multiple studies to obtain an understanding of (a) the magnitude of an intervention's effect across studies and (b) whether the magnitude of effect varies according to certain characteristics of the studies. The results from the meta-analysis can provide a useful basis for analyzing the benefits and costs of a program and can inform program adoption decisions.

Below, we describe the process that we followed to conduct our meta-analysis of arts education studies, focusing, in particular, on the following issues:

- ▶ The criteria for including findings from studies in the meta-analysis
- ▶ The method for converting study-reported findings into standardized effect sizes
- ▶ The method for aggregating effects across studies and exploring variations by art type and outcome domain

Eligibility of Studies for Inclusion in the Meta-Analysis

As our WWC-certified reviewers examined each study to determine whether its findings provide evidence in ESSA's Tiers I–III, they focused on four features: (a) the study's research design, (b) whether the study showed statistically significant effects of arts education, (c) whether the study was conducted with more than 349 students, and (d) whether the study was conducted in

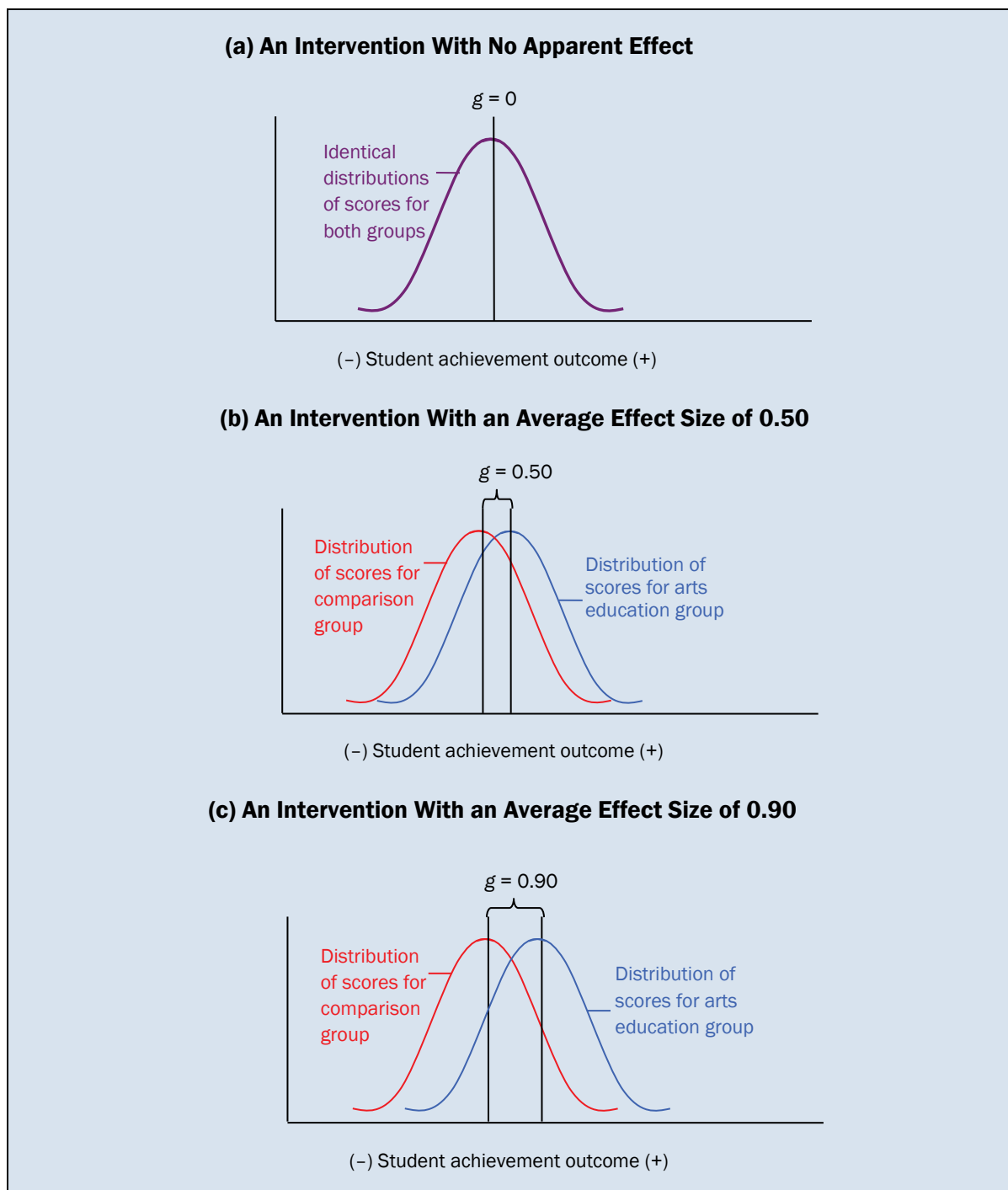
more than one site (see Table A.1 in the prior section). To be eligible for the meta-analysis, studies only had to meet the research design criterion and contain effect estimates that could be converted to effect sizes. Thus, the meta-analysis includes findings that study authors reported as not statistically significant, findings from studies involving fewer than 350 students, and studies conducted in a single site.

The meta-analysis does not include findings from studies that did not use research designs necessary for meeting the criteria for Tiers I–III. This design requirement excluded pre-post studies common in action research, correlational studies that failed to control for potential confounds statistically, descriptive studies that did not use a comparison group, and group design studies in which the intervention and comparison groups were not equivalent at baseline. Regarding the latter group of studies (i.e., quasi-experimental studies with nonequivalent groups at baseline), an exception was made if the variable that distinguishes the groups at baseline was controlled for in the impact analysis (i.e., met the definition of Tier III).

Calculation of Effect Sizes

Hedges' g Statistic. Following the practice of the WWC, our review team used Hedges' g statistic as the effect size metric. Conceptually, Hedges' g reflects the difference between the intervention group's average score and the comparison group's average score scaled according to the pooled standard deviation of the groups (see examples in Figure A.2). For example, a g statistic of +0.50 indicates that the average outcomes for the two groups differ by half of a standard deviation.

Figure A.2. Examples of Possible Effects of an Arts Education Intervention on Student Academic Achievement



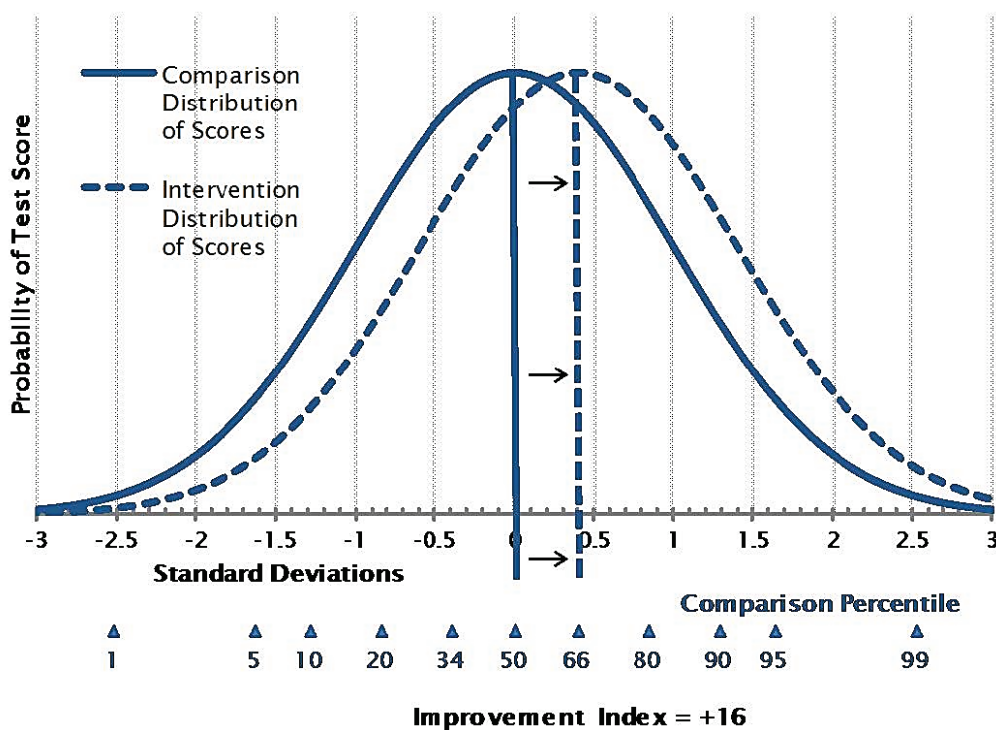
Our WWC-certified reviewers used WWC Study Review Guide to record study features and convert test statistics found in the study reports into Hedges' g . The Excel-based WWC Study Review Guide is macro-enabled and contains functions to determine whether attrition in a randomized controlled trial was high or low and whether study groups in a high-attrition RCT or a quasi-

experimental study were equivalent at baseline. It can also convert different types of test statistics into Hedges' g .

The one limitation of Hedges' g is that it fails to convey the magnitude of an effect in a language that educators, school and district administrators, and policymakers can easily understand. One benchmark that the WWC has adopted is the effect size of 0.25, which is the threshold for labeling an effect "substantively important."

Improvement Index. Another practitioner- and policymaker-friendly way to convey the magnitude of an intervention's effect is to report the improvement index associated with the effect. The improvement index reflects the percentile point gain that an average student (i.e., the student at the 50th percentile) in the comparison group would experience had he or she been exposed to the intervention. For example, an effect size of 0.40 corresponds to an improvement index of 16, which indicates that the average student in the comparison group would have scored 16 percentile points higher had he or she received the intervention. See Figure A.3 for an illustration.

Figure A.3. Computing and Interpreting the Improvement Index



Source: Reprinted with permission from WWC, 2014, p. 15 (Figure IV.1).

Aggregation of Effect Sizes Across Studies

After converting relevant study findings to effect sizes, we calculated the standard error associated with each effect size based on the sample size of each group and the effect size (see equation in Borenstein, Hedges, Higgins, & Rothstein, 2009, pp. 226–227).²² Conceptually, the standard error indicates the amount of uncertainty surrounding an effect size estimated based on a particular sample. As the sample size increases, the standard error of the estimate decreases.

We then calculated a pooled effect size using meta-analytic methods that produce a precision-weighted estimate and account for dependence among effect size estimates (e.g., effect size estimates from the same study tend to be correlated; see Tanner-Smith et al., 2016). Our inclusion criteria produced a total of 77 effect sizes across 20 studies; many studies contributed more than one effect size. We estimated a standard random-effects model using a robust variance estimation method (Hedges et al., 2010; Tanner-Smith et al., 2016), which accounts for both the varying degrees of precision of estimates across studies and the nonindependence of effect sizes within studies. We used *robumeta*, an R-package for robust variance estimation, to perform the meta-analysis. The weighted average effect size across all studies was then presented along with its 95% confidence interval. This interval is represented by upper and lower bounds. The “true effect” for the population was estimated (with 95% certainty) to be among the values within the interval bounds. In addition to estimating the overall average effect, we conducted subgroup analyses to estimate the average effect by art type and outcome domain and to test whether the average effect varies significantly by art type and outcome domain.

²² Per best practices in meta-analysis, an adjustment is made to the standard errors that come from studies for which the unit of assignment did not match the unit of analysis (such as when schools are randomly assigned to different study groups, and yet the analysis was done at the student level and failed to account for the clustered data structure). Hedges (2009) provided an adjustment that can be made to an effect’s standard error so that it approximates the standard error that would be obtained had the analysis properly accounted for clustering.

Appendix B. Evidence-Based Criteria for Grant Programs Awarded by the U.S. Department of Education

Definitions that apply to all Department programs, 34 C.F.R. § 77.1 (U.S. Department of Education, 2017).

Demonstrates a rationale means a key project component included in the project's logic model is informed by research or evaluation findings that suggest the project component is likely to improve relevant outcomes.

Logic model (also referred to as a theory of action) means a framework that identifies key project components of the proposed project (i.e., the active "ingredients" that are hypothesized to be critical to achieving the relevant outcomes) and describes the theoretical and operational relationships among the key project components and relevant outcomes.

Moderate evidence means that there is evidence of effectiveness of a key project component in improving a relevant outcome for a sample that overlaps with the populations or settings proposed to receive that component, based on a relevant finding from one of the following:

- (i) A practice guide prepared by the WWC using version 2.1 or 3.0 of the WWC Handbook reporting a "strong evidence base" or "moderate evidence base" for the corresponding practice guide recommendation;
- (ii) An intervention report prepared by the WWC using version 2.1 or 3.0 of the WWC Handbook reporting a "positive effect" or "potentially positive effect" on a relevant outcome based on a "medium to large" extent of evidence, with no reporting of a "negative effect" or "potentially negative effect" on a relevant outcome; or
- (iii) A single experimental study or quasi-experimental design study reviewed and reported by the WWC using version 2.1 or 3.0 of the WWC Handbook, or otherwise assessed by the Department using version 3.0 of the WWC Handbook, as appropriate, and that—
 - (A) Meets WWC standards with or without reservations;
 - (B) Includes at least one statistically significant and positive (i.e., favorable) effect on a relevant outcome;
 - (C) Includes no overriding statistically significant and negative effects on relevant outcomes reported in the study or in a corresponding WWC intervention report prepared under version 2.1 or 3.0 of the WWC Handbook; and
 - (D) Is based on a sample from more than one site (e.g., State, county, city, school district, or postsecondary campus) and includes at least 350 students or other individuals across sites. Multiple studies of the same project component that each meet requirements in paragraphs (iii)(A), (B), and (C) of this definition may together satisfy this requirement.

Promising evidence means that there is evidence of the effectiveness of a key project component in improving a relevant outcome, based on a relevant finding from one of the following:

- (i) A practice guide prepared by WWC reporting a “strong evidence base” or “moderate evidence base” for the corresponding practice guide recommendation;
- (ii) An intervention report prepared by the WWC reporting a “positive effect” or “potentially positive effect” on a relevant outcome with no reporting of a “negative effect” or “potentially negative effect” on a relevant outcome; or
- (iii) A single study assessed by the Department, as appropriate, that—
 - (A) Is an experimental study, a quasi-experimental design study, or a well-designed and well-implemented correlational study with statistical controls for selection bias (e.g., a study using regression methods to account for differences between a treatment group and a comparison group); and
 - (B) Includes at least one statistically significant and positive (i.e., favorable) effect on a relevant outcome.

Strong evidence means that there is evidence of the effectiveness of a key project component in improving a relevant outcome for a sample that overlaps with the populations and settings proposed to receive that component, based on a relevant finding from one of the following:

- (i) A practice guide prepared by the WWC using version 2.1 or 3.0 of the WWC Handbook reporting a “strong evidence base” for the corresponding practice guide recommendation;
- (ii) An intervention report prepared by the WWC using version 2.1 or 3.0 of the WWC Handbook reporting a “positive effect” on a relevant outcome based on a “medium to large” extent of evidence, with no reporting of a “negative effect” or “potentially negative effect” on a relevant outcome; or
- (iii) A single experimental study reviewed and reported by the WWC using version 2.1 or 3.0 of the WWC Handbook, or otherwise assessed by the Department using version 3.0 of the WWC Handbook, as appropriate, and that—
 - (A) Meets WWC standards without reservations;
 - (B) Includes at least one statistically significant and positive (i.e., favorable) effect on a relevant outcome;
 - (C) Includes no overriding statistically significant and negative effects on relevant outcomes reported in the study or in a corresponding WWC intervention report prepared under version 2.1 or 3.0 of the WWC Handbook; and
- (iv) Is based on a sample from more than one site (e.g., State, county, city, school district, or postsecondary campus) and includes at least 350 students or other individuals across sites. Multiple studies of the same project component that each meet requirements in paragraphs (iii)(A), (B), and (C) of this definition may together satisfy this requirement.



AMERICAN INSTITUTES FOR RESEARCH®