

Exploring Links Between Arts Education and Academic Outcomes in the International Baccalaureate

Kenneth Elpus, Ph.D. David Miller



ACKNOWLEDGEMENTS

The research reported here was supported by the Institute of Education Sciences, U.S Department of Education, through Grant R305A180062 to the University of Maryland.

The opinions expressed here are those of the authors and do not represent the view of the Institute or the U.S. Department of Education.





ACKNOWLEDGEMENTS

The research reported was supported by the Maryland Longitudinal Data System (MLDS) Center. We appreciate the feedback received from the MLDS Center and its stakeholder partners.

All opinions are the authors' and do not represent the opinion of the MLDS Center or its partner agencies.





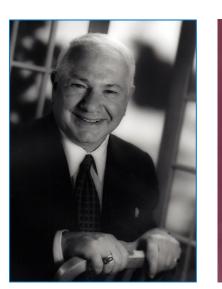
Arts Education in Maryland

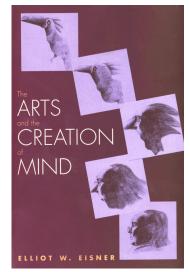
COMAR 13a.04.16.01 requires that each local school system in Maryland

- Provide arts education in dance, media arts, music, theatre, and visual art for all students in PreK through 8th grade
 - PreK through 5th graders must have experiences in all arts forms
 - Students in 6th through 8th may specialize in one or more arts disciplines
- Provide an instructional program allowing all 9th through 12th grade students to meet graduation requirements by choosing from among courses in dance, media arts, music, theatre, and visual art

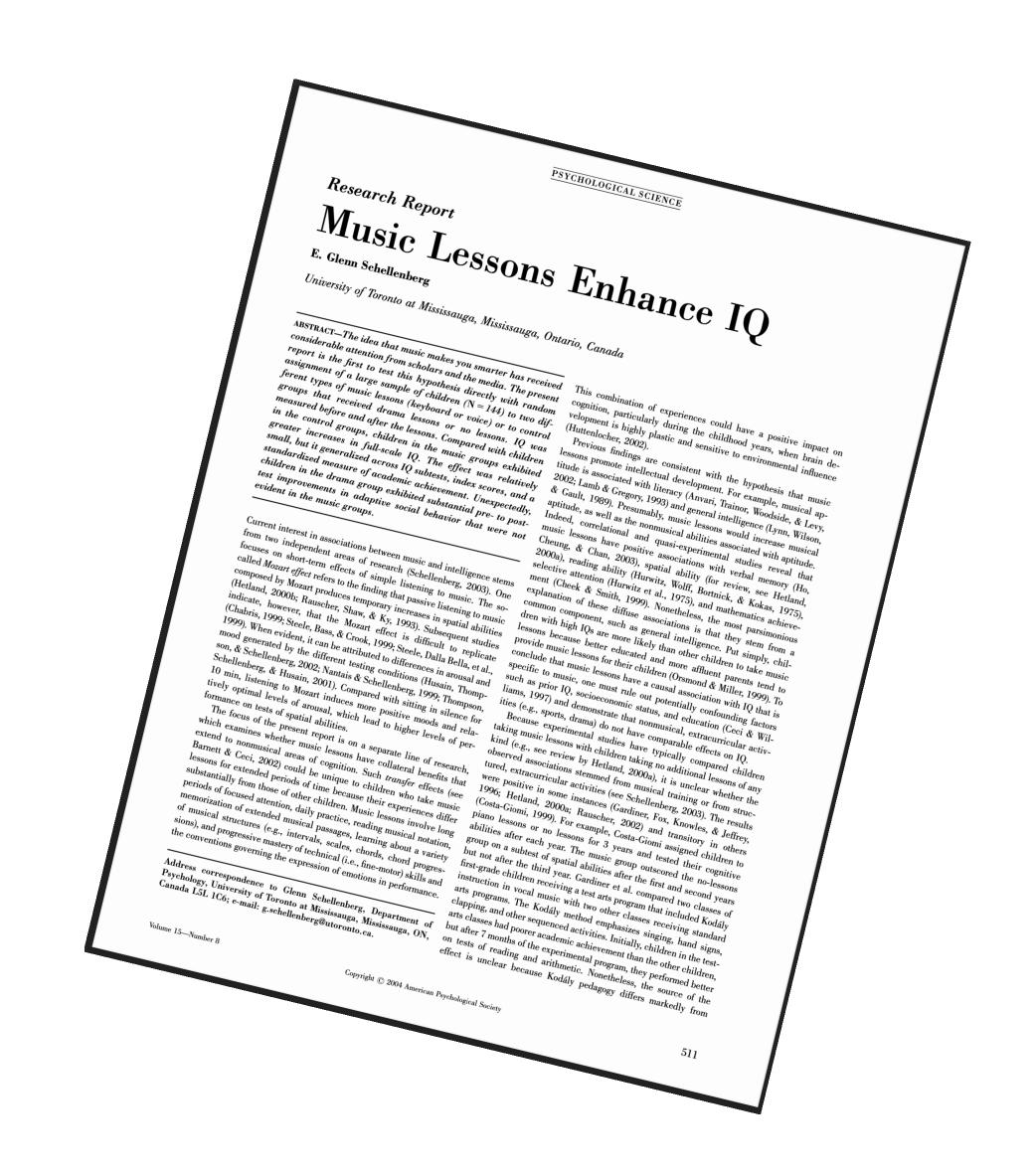
Aims of Arts Education

"The tasks the arts put forward—such as noticing subtleties among qualitative relationships, conceiving of imaginative possibilities, interpreting the metaphorical meanings the work displays, exploiting unanticipated opportunities in the course of one's work—require complex cognitive modes of thought."









SCIENTIFIC CORRESPONDENCE

Music and spatial task performance

abstract operations such as mathematical or spatial reasoning. We performed an or spatial reasoning. We performed an or spatial reasoning. We performed an

Testing procedure. In the music condition, the subject tailed). Pulse rates were tailed). Pulse rates were tailed to 10 min of the Mozart piece. The relaxation tailed to 10 min of the Mozart piece and 10 min of relaxation to 10 min of the subject to listent a 10 min of relaxation to 10 min of the subject to 10 min of relaxation to 10 min of the subject to 10 min of t listened to 10 min of the Mozart piece. The relaxation condition required the subject to listen to 10 min of relaxation taken before and after taken before and instructions designed to lower blood pressure. The silence condition required the subject to sit in silence for 10 min. One condition required the subject to sit in silence for 10 min. One of three abstract reasoning tests taken from the Stanford discovered time of pulse. condition required the subject to sit in silence for 10 min. One of three abstract reasoning tests taken from the Stanford dition and time of pulse dition and time of puls of three abstract reasoning tests taken from the Stanford— Binet intelligence scale⁴ was given after each of the listening ginet intelligence scale⁴ was given after each of the listening measure) repeated measure) repeated measure) repeated measure) repeated measure) revealed the down regulation of c-fos expression dition and time of pulse during muscle cell differentiation may during muscle of the helix-loop-result from the binding of the he conditions. The abstract/spatial reasoning tasks consisted of a pattern analysis test, a multiple-choice matrices test and a multiple-choice paper-folding and cutting test For our multiple-choice paper-folding and cutting test For a pattern analysis test, a multiple-choice matrices test and a multiple-choice paper-folding and cutting test. For our sample, these three tasks correlated at the 0.01 level of sample. multiple-choice paper-folding and cutting test. For our sample, these three tasks correlated at the 0.01 level of sample, these three tasks correlated at the mass equal significance. We were thus able to treat them as equal excluding arousal as an exclusion arousal arou sample, these three tasks correlated at the 0.01 level of significance. We were thus able to treat them as equal measures of abstract reasoning ability.

equivalents were calculated by first multiplying each SAS by 3 (the number of subtests required by the Stanford-Binet for calculating IOs). We then used their area score conversion calculating IOs). We then used their area score conversion to a local score conversion calculating IOs). We then used their area score conversion to a local score conversion calculating IOs). (the number of subtests required by the Stanford-Binet for calculating IQs). We then used their area score conversion table, designed to have a mean of 100 and a standard table, designed to have a mean of 100 and a standard table. calculating IQS). We then used their area score conversion table, designed to have a mean of 100 and a standard table, designed to have a mean of 100 and a standard jects were engaged in each 1 deviation of 16, to obtain SAS IQ equivalents.

given three sets of standard IQ spatial reasoning tasks; each task was preceded by 10 minutes of (1) listening to Mozart's sonata for two pianos in D major, K488; (2) listening to a relaxation tape; or (3) silence. Performance was improved for those tasks immediately following the first condition compared to the second two.

Thirty-six college students participated tion, the student's spatial reasoning skills other compositions and musical styles other compositions. were tested using the Stanford-Binet intelligence scale⁴. The mean standard age scores (SAS) for the three listening conditions are shown in the figure. The music condition yielded a mean SAS of 57.56; the mean SAS for the relaxation condition was 54,61 and the mean score for the silent condition was 54.00. To assess of these scores, we 'translated' them to word. NATURE · VOL 365 · 14 OCTOBER 1993

and anecdotal' relationships between music cognition and other 'higher brain functions' but no causal relationship the music condition were participating in the music condition were participating to compare these two grounds. functions', but no causal relationship has been demonstrated between music specified and specified a nunctions, but no causal relationship has been demonstrated between music the other two conditions. A one-factor the other two conditions. A one-factor the other two conditions conditions are conditions and conditions are conditions. the other two conditions. A one-factor cognition and cognitions pertaining to abstract operations such as mathematical abstract operations.

stract/spatial reasoning tests after listening to Mozart than after listening to either the relaxation tape or to nothing $(F_{2,35} = 7.08;$ P = 0.002). The music condition differed significantly from both the relaxation and the silence conditions (Scheffe's t = 3.41, P =0.002; t = 3.67, P = 0.0008,respectively). The relaxation and silence

determine the presence of a decay constant. It would also be interesting to vary the listening time to optimize the enhancing effect, and to examine whether other measures of general intelligence (verbal short-term memory) would be similarly facilitated. Because we used only one musical sample of one composer, various

> Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in Nature. In any case, priority will be given to letters of fewer than 500

music lacking complexity or which is repetitive may interfere with, rather than enhance, abstract reasoning. Also, as musicians may process music in a different

Center for the Neurobiology of Learning and Memory, University of California, Irvine, California 92717, USA

MyoD and c-fos

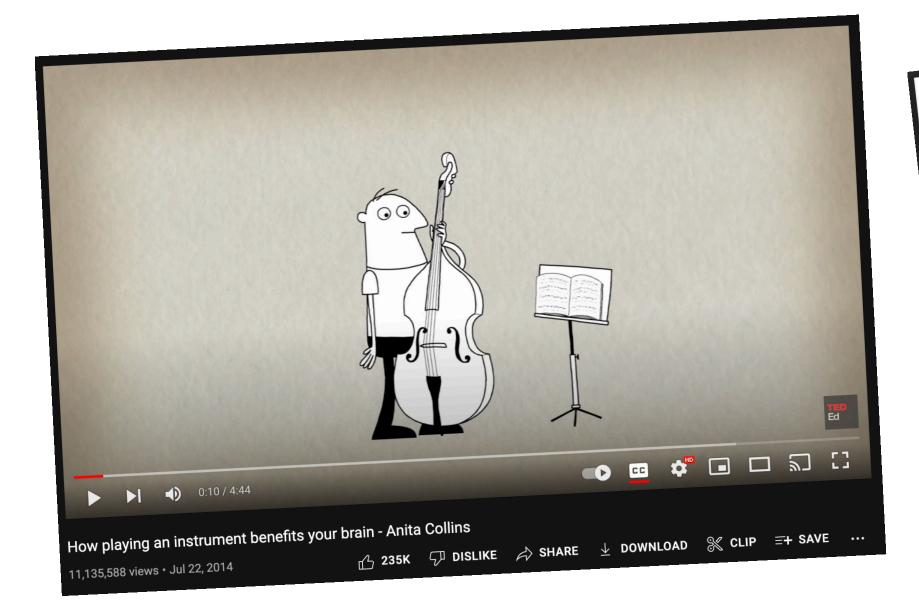
measures of abstract reasoning ability.

Scoring. Raw scores were calculated by subtracting the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number of items failed from the number of items failed from the highest item number of items failed from the highest item number of items failed from the number of items failed from the highest item number of items failed from the highest item number of items failed from the number Scoring. Raw scores were calculated by subtracting the number of items failed from the highest item number administered. These were then converted to SAS using the Stanfordnered. These were then converted to SAS using the Stanfordnered standard scores any experimenter effect. the myogenin/E12-SRE complex was 10^{-8} - 10^{-9} M⁻¹ by comparing the relative spatial task. Inclusion of a spatial task. Inclusion of a different genes to compete for binding in different gene E12 for the SRE could result from differences in nucleotide identities at the internal dinucleotide and flanking sequences between the c-fos E-box and the consensus HLH binding site^{3,4} measures or general intelligence (verbal reasoning, quantitative reasoning and short-term memory) would be similarly Collectively, these data indicate that it is unlikely that HLH proteins alone can significantly compete with SRF for binding to DNA in vivo.

It has been reported in other studies that the c-fos SRE is either equally active in muscle and non-muscle cells6, or that it activates muscle-specific expression when situated upstream from a minimal promoter. Further, a comparison of the

Music Lessons Enhance IQ

© 1993 Nature Publishing Group



Music and spatial task performance

functions', but no causal relationship has been demonstrated between music specified and specified a nunctions, but no causal relationship has been demonstrated between music the other two conditions. A one-factor the other two conditions. A one-factor the other two conditions conditions are conditions and conditions are conditions. the other two conditions. A one-factor cognition and cognitions pertaining to abstract operations such as mathematical abstract operations. cognition and cognitions pertaining to abstract operations such as mathematical or spatial reasoning. We performed an or spatial reasoning.

Testing procedure. In the music condition, the subject tailed to 10 min of the Mozart piece. The relaxation tailed to 10 min of the subject to listen to 10 min of relaxation and the subject to listen to 10 min of relaxation to 10 min of the 10 min of relaxation to 10 min of relaxation to 10 min of the 10 min of relaxation to 10 min of the 10 min of relaxation to 10 min of the 10 min of th listened to 10 min of the Mozart piece. The relaxation condition required the subject to listen to 10 min of relaxation taken before and after taken before and instructions designed to lower blood pressure. The silence condition required the subject to sit in silence for 10 min. One condition required the subject to sit in silence for three abstract reasoning tests taken from the Stanford three abstract reasoning tests three

equivalents were calculated by first multiplying each SAS by 3 (the number of subtests required by the Stanford-Binet for calculating IOs). We then used their area score conversion calculating IOs). We then used their area score conversion to the subtest required by the Stanford-Binet for calculating IOs). We then used their area score conversion to the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the Stanford-Binet for part of the subtest required by the subtest (the number of subtests required by the Stanford-Binet for calculating IQs). We then used their area score conversion table, designed to have a mean of 100 and a standard table, designed to have a mean of 100 and a standard table. with a mean set at 50 and a standard deviation of 8. IQ equivalents were calculated by first multiplying each SAS by 3 (the number of subtests required by the Stanford Block for

given three sets of standard IQ spatial

Thirty-six college students participated tion, the student's spatial reasoning skills other compositions and musical styles other compositions. were tested using the Stanford-Binet intelligence scale⁴. The mean standard age scores (SAS) for the three listening conditions are shown in the figure. The music condition yielded a mean SAS of 57.56; the mean SAS for the relaxation condition condition was 54.00. To assess

music lacking complexity or which is repetitive may interfere with, rather than enhance, abstract reasoning. Also, as musicians may process music in a different

Center for the Neurobiology of Learning and Memory, University of California, Irvine, California 92717, USA

conditions did not differ (t = 0.795; P = 0.432, two-toiled) Pulse rates were

measures of abstract reasoning ability.

Scoring. Raw scores were calculated by subtracting the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number administrates and the number of items failed from the highest item number of items failed from the number of items failed from the highest item number of items failed from the number of item Scoring. Raw scores were calculated by subtracting the number of items failed from the highest item number administered. These were then converted to SAS using the Stanfordnered. These were then converted to SAS using the Stanfordnered standard scores any experimenter effect. the myogenin/E12-SRE complex was 10^{-8} - 10^{-9} M⁻¹ by comparing the relative spatial task. Inclusion of a spatial task. Inclusion of a different genes to compete for binding in different gene E12 for the SRE could result from differences in nucleotide identities at the determine the presence of a decay coninternal dinucleotide and flanking sestant. It would also be interesting to vary quences between the c-fos E-box and the listening time to optimize the enhancthe consensus HLH binding site^{3,4} measures or general intelligence (verbal reasoning, quantitative reasoning and short-term memory) would be similarly ing effect, and to examine whether other measures of general intelligence (verbal facilitated. Because we used only one Collectively, these data indicate that it musical sample of one composer, various is unlikely that HLH proteins alone can significantly compete with SRF for

binding to DNA in vivo. It has been reported in other studies that the c-fos SRE is either equally active in muscle and non-muscle cells⁶, or that it activates muscle-specific expression when situated upstream from a minimal promoter. Further, a comparison of the

Music Lessons Enhance IQ

SCIENTIFIC CORRESPONDENCE

and anecdotal' relationships between music cognition and other 'higher brain functions', but no causal relationship the music condition were participating in the music condition were participating to compare these two around the music conditions are provided to the music condition were participating to compare these two around the music conditions are provided to the music condition were participating to compare these two around the music condition were participating to compare these two around the music condition were participating to compare the music condition were provided to the music condition were participating to compare the music condition were participating to the music condition were participating to the music condition were participating to

or spatial reasoning. We performed an

stract/spatial reasoning tests after listening to Mozart than after listening to either the relaxation tape P = 0.002). The music condition differed significantly from both the relaxation and the silence conditions (Scheffe's t = 3.41, P =0.002; t = 3.67, P = 0.0008,

condition required the subject to sit in silence for 10 min. One of three abstract reasoning tests taken from the Stanford dition and time of pulse dition and time of puls of three abstract reasoning tests taken from the Stanford— Binet intelligence scale⁴ was given after each of the listening generated intelligence scale⁴ was given after each of the listening measure) repeated measure repeated measure) repeated measure repeate conditions. The abstract/spatial reasoning tasks consisted of a pattern analysis test, a multiple-choice matrices test and a multiple-choice paper-folding and cutting test For our multiple-choice paper-folding and cutting test For multiple-choice paper-folding and cutting test. For our sample, these three tasks correlated at the 0.01 level of sample, these three tasks correlated at the mass equal significance. We were thus able to treat them as equal excluding arousal as an exclusion arousal arou sample, these three tasks correlated at the 0.01 level of significance. We were thus able to treat them as equal measures of abstract reasoning ability.

calculating IQS). We then used their area score conversion table, designed to have a mean of 100 and a standard deviation of 16, to obtain SAS IQ equivalents.

reasoning tasks; each task was preceded by 10 minutes of (1) listening to Mozart's sonata for two pianos in D major, K488; (2) listening to a relaxation tape; or (3) silence. Performance was improved for those tasks immediately following the first condition compared to the second two.

was 54,61 and the mean score for the silent of these scores, we 'translated' them to word. NATURE · VOL 365 · 14 OCTOBER 1993

© 1993 Nature Publishing Group

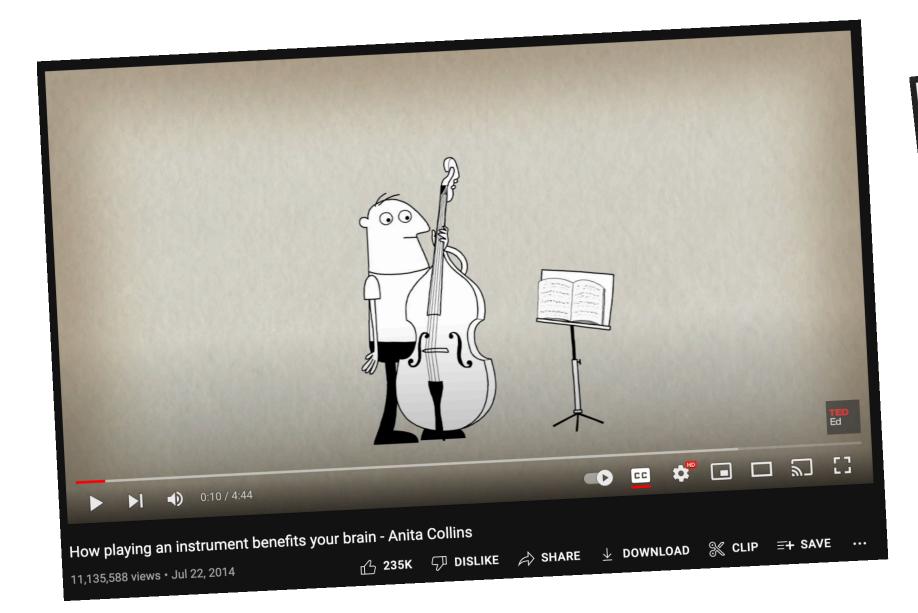
Scientific Correspondence is intended to

provide a forum in which readers may raise points of a scientific character.

They need not arise out of anything

published in Nature. In any case, priority

will be given to letters of fewer than 500





Power of Art: Can painting improve your grades?

and anecdotar relationships between music cognition and other 'higher brain functions' but to could relationship the other two conditions. A one-ractor cognition and cognitions pertaining to abstract operations such as mathematical abstract operations such as mathematical analysis of vortices (ANOVA) area. abstract operations such as mathematical or spatial reasoning. We performed an or spatial reasoning. We performed an

listened to 10 min of the Mozart piece. The relaxation tailed). Pulse rates were condition required the subject to listen to 10 min of relaxation taken before and after

given three sets of standard IQ spatial condition compared to the second two.

Thirty-six college students participated in all three listening conditions. Immediately following each listening condition, the student's spatial reasoning skills were tested using the Stanford-Binet inscores (SAS) for the three listening condicondition yielded a mean SAS of 57.56; the mean SAS for the relaxation condition condition was 54.00. To assess

petitive may interfere with, rather than enhance, abstract reasoning. Also, as musicians may process music in a different way from non-musicians, it would be

interesting to compare these two groups. Gordon L. Shaw*

after listening to Mozart than after listening to

dition differed significantly and the silence conditions 0.002; t = 3.67, P = 0.0008,

Center for the Neurobiology of Learning and Memory, University of California, Irvine, California 92717, USA stract/spatial reasoning tests

MyoD and c-fos

dition and time of pulse the down regulation of c-fos expression measure) repeated meaduring muscle cell differentiation may motif, or E-box, that occurs within the obvious cause. We found no thereby excluding the binding of the measures of abstract reasoning ability.

Scoring. Raw scores were calculated by subtracting the number of items failed from the highest item number administrated and the number of items failed from the highest item number administrated and the interaction or task and the interaction of the number of items failed from the highest item number administrated and the interaction of the interacti Scoring. Raw scores were calculated by subtracting the number of items failed from the highest item number administrated. These were then converted to SAS using the Stanford-target. These were then converted to SAS using the Stanford-target. the music condition is temmated that the dissociation constant for the myogenin/E12-SRE complex was 10^{-8} - 10^{-9} M⁻¹ by comparing the relative spatial task. Inclusion of a spatial task. Inclusion of a delay period (as a variable) between the delay period (as a variable) between the delay period (as a variable) and the testing delay period (as a variable) between the electrophoratic mobility-chift accounts the electrophoratic mobility-chift accounts. E12 for the SRE could result from differences in nucleotide identities at the determine the presence of a decay coninternal dinucleotide and flanking sestant. It would also be interesting to vary quences between the c-fos E-box and the listening time to optimize the enhancthe consensus HLH binding site^{3,4} contrast, the dissociation constants for the SRE-SRF or the SRE-SRF/p62^{TCF} ing effect, and to examine whether other measures of general intelligence (verbal short-term memory) would be similarly 5×10^{-10} and $\leq 10^{-11} \text{ M}^{-1}$, respectively⁵. facilitated. Because we used only one Collectively, these data indicate that it musical sample of one composer, various is unlikely that HLH proteins alone other compositions and musical styles

can significantly compete with SRF for binding to DNA in vivo. It has been reported in other studies that the c-fos SRE is either equally active in muscle and non-muscle cells6, or that it activates muscle-specific expression when situated upstream from a minimal promoter. Further, a comparison of the

Music Lessons Enhance IQ

SCIENTIFIC CORRESPONDENCE

Music and spatial task performance respectively. Thus, the IQs of subjects

condition required the subject to listen to 10 min of relaxation taken before and after each listening condition. A instructions designed to lower blood pressure. The shellow each listening condition. A two-factor (listening condition required the subject to sit in silence from the Stanfordconditions. The abstract/spatial reasoning tasks consisted of a pattern analysis test, a multiple-choice matrices test and a multiple-choice paper-folding and cutting test For our multiple-choice paper-folding and cutting test For multiple-choice paper-folding and cutting test. For our no interaction or main sample, these three tasks correlated at the 0.01 level of effects for pulse, thereby eignificance. We were this able to treat them as equal sample, these three tasks correlated at the U.UI level of effects for pulse, thereby significance. We were thus able to treat them as equal excluding arousal as an excluding arousal as a constant are excluding arousal as a constant arousal arousa

tered. These were then converted to SAS using the Stanford equivalents were calculated by first multiplying each SAS by 3 (the number of subtests required by the Stanford–Binet for calculating lOs). We then used their area score conversion calculating lOs). We then used their area score conversion (the number of subtests required by the Stanford-Binet for calculating IQs). We then used their area score conversion table. designed to have a mean of 100 and a standard table. designed to have a mean of 100 and a standard table. calculating IQS). We then used their area score conversion table, designed to have a mean of 100 and a standard table, designed to have a mean of 100 and a standard during which subjects were engaged in each deviation of 16, to obtain SAS IQ equivalents.

reasoning tasks; each task was preceded by 10 minutes of (1) listening to Mozart's sonata for two pianos in D major, K488; (2) listening to a relaxation tape; or (3) silence. Performance was improved for those tasks immediately following the first

The mean standard age was 54.61 and the mean score for the silent of these scores, we 'translated' them to NATURE · VOL 365 · 14 OCTOBER 1993

© 1993 Nature Publishing Group

Scientific Correspondence is intended to

raise points of a scientific character

They need not arise out of anything

published in Nature. In any case, priority

will be given to letters of fewer than 500



REQUEST FOR APPLICATIONS

Education Research Grants CFDA Number: 84.305A

> Arts in Education

Program Officers: Dr. James Benson (202-245-8333; <u>James.Benson@ed.gov</u>)

Dr. Erin Higgins (202-245-6541; Erin.Higgins@ed.gov)

a) Purpose

The Arts in Education special topic supports research to understand the implementation and effects of arts programs and policies at the K-12 level in order to improve the education outcomes of students. Research connecting student participation in the arts to academic outcomes and social/behavioral competencies has the potential to inform contemporary policy debates regarding the benefits of arts programming in schools. Advocates of the arts have long argued for their inclusion in schools, for their general benefits, such as improved innovation, creativity, and communication (Winner, Goldstein, and Vincent-Lancrin, 2013), as well as for their perceived positive effects on literacy (Walker, Tabone, and Weltsek 2011; Podlozny 2000), math achievement (Courey, Balogh, and Siker 2012; Kinney and Forsythe 2005), critical thinking (Montgomerie and Ferguson 1999), and engagement in school (Smithrim and Upitis 2005). In addition, there is some evidence from cognitive psychology and neuroscience suggesting a relationship between participation in the arts and improved cognitive and neural processing (e.g., Catterall, 2002; Tierney, Krizman, and Kraus 2015; Kraus, Hornickel, Strait, Slater, and Thompson, 2014).

States and school districts often feel the need to make tradeoffs between instruction in core subjects (e.g., math, reading) and instruction in the arts, in part because of the emphasis on testing in core subjects as well as because of budgetary pressure. Given the potential of the arts to contribute positively to students' success in school, new research is needed to rigorously assess the effect of arts participation on education outcomes, including a close look at potential mediators of any effects, the types of outcomes impacted, and the conditions under which these relationships hold.



REQUEST FOR APPLICATIONS

Education Research Grants CFDA Number: 84.305A

> Arts in Education

Program Officers: Dr. James Benson (202-245-8333; <u>James.Benson@ed.gov</u>)

Dr. Erin Higgins (202-245-6541; Erin.Higgins@ed.gov)

a) Purpose

The Arts in Education special topic supports research to understand the implementation and effects of arts programs and policies at the K-12 level in order to improve the education outcomes of students. Research connecting student participation in the arts to academic outcomes and social/behavioral competencies has the potential to inform contemporary policy debates regarding the benefits of arts programming in schools. Advocates of the arts have long argued for their inclusion in schools, for their general benefits, such as improved innovation, creativity, and communication (Winner, Goldstein, and Vincent-Lancrin, 2013), as well as for their perceived positive effects on literacy (Walker, Tabone, and Weltsek 2011; Podlozny 2000), math achievement (Courey, Balogh, and Siker 2012; Kinney and Forsythe 2005), critical thinking (Montgomerie and Ferguson 1999), and engagement in school (Smithrim and Upitis 2005). In addition, there is some evidence from cognitive psychology and neuroscience suggesting a relationship between participation in the arts and improved cognitive and neural processing (e.g., Catterall, 2002; Tierney, Krizman, and Kraus 2015; Kraus, Hornickel, Strait, Slater, and Thompson, 2014).

States and school districts often feel the need to make tradeoffs between instruction in core subjects (e.g., math, reading) and instruction in the arts, in part because of the emphasis on testing in core subjects as well as because of budgetary pressure. Given the potential of the arts to contribute positively to students' success in school, new research is needed to rigorously assess the effect of arts participation on education outcomes, including a close look at potential mediators of any effects, the types of outcomes impacted, and the conditions under which these relationships hold.

Are There Actually Arguments Made Against the Arts?

Are There Actually Arguments Made Against the Arts?

Intensive Math Instruction and **Educational Attainment** Long-Run Impacts of Double-Dose Algebra

Kalena E. Cortes Joshua S. Goodman Takako Nomi

We study an intensive math instruction policy that assigned low-skilled ninth graders to an algebra course that doubled instructional time, altered peer composition and emphasized problem solving skills. A regression peer composition and emphasized problem solving skills. A regression discontinuity design shows substantial positive impacts of double-dose algebra on credits earned, test scores, high school graduation, and college enrollment rates. Test score effects underpredict attainment effects, highlighting the importance of long-run evaluation of such a policy. Perhaps because the intervention focused on verbal exposition of mathematical concepts, the impact was largest for students with below-average reading skills, emphasizing the need to target interventions toward appropriately skilled students.

Kalena E. Cortes is an assistant professor of public policy at Texas A&M University and a Faculty Research Fellow at the National Burgay of Franchic Passarch Lookus S. Goodman is an assistant professor. Kalena E. Cortes is an assistant professor of public policy at Texas A&M University and a Faculty Research Fellow at the National Bureau of Economic Research. Joshua S. Goodman is an assistant professor of public policy at Harvard University and a Faculty Passarch Fallow at the National Bureau of Faculty Passarch Fallow at the National Bureau o search renow at the National Bureau of Economic Research. Joshua S. Goodman is an assistant professor of public policy at Harvard University and a Faculty Research Fellow at the National Bureau of Economic Research. Takaba Nami is an assistant professor of advertion at St. Lovie University. Some of the Research Takaba Nami is an assistant professor of advertion at St. Lovie University. of public policy at Harvara University and a Faculty Research Fellow at the National Bureau of Economic Research. Takako Nomi is an assistant professor of education at St. Louis University. Some of this research кезеатся. такако мот is an assistant projessor of eaucation at St. Louis University. Some of this rest was conducted while Cortes was a Visiting Scholar at Stanford's Graduate School of Education. The was conauctea white Cortes was a visiting Schools for sharing their data and to Sue Sporte, Director of Reauthors are grateful to the Chicago Public Schools for sharing their data and to Sue Sporte, Director of Research Operations Consortium on Chicago School Personal (CCSP) for facilitating this charing They also search Operations Consortium on Chicago School Personal (CCSP) for facilitating this charing They also authors are grateful to the Unicago Public Schools for sharing their data and to Sue Sporte, Director of Kesearch Operations, Consortium on Chicago School Research (CCSR) for facilitating this sharing. They also search Operations, Consortium on Chicago School Research (CCSR) for facilitating this sharing. Tools Under Town Long Laffrey Vishil Low Toylor Look Walnut School Research (CCSR) for facilitating this sharing. They also search Operations, Consortium on Chicago School Research (CCSR) for facilitating this sharing. They also search Operations, Consortium on Chicago School Research (CCSR) for facilitating this sharing. They also search Operations, Consortium on Chicago School Research (CCSR) for facilitating this sharing. They also search Operations, Consortium on Chicago School Research (CCSR) for facilitating this sharing. search Operations, Consortium on Chicago School Research (CCSR) for Jacillating this sharing. They also thank for helpful comments Richard Murnane, Bridget Terry Long, Jeffrey Kubik, Lori Taylor, Jacob Vigdor, thank for helpful comments Richard Murnane, Bridget Terry Long, Jeffrey Rubik, Lori Taylor, Jacob Vigdor, thank for helpful comments Richard Murnane, Bridget Terry Long, Jeffrey Rubik, Lori Taylor, Jacob Vigdor, as well as seminar and conference participation. thank for neight comments Richard Murnane, Bridget Terry Long, Jeffrey Kubik, Lori Taylor, Jacob Vigdor, Caroline Hoxby, Martin West, Kevin Stange, and Nora Gordon, as well as seminar and conference participated the State of Texas Education Research. Caroune Hoxby, Martin West, Kevin Stange, and Nora Gordon, as well as seminar and conference participants at Harvard's Program on Education Policy and Governance, the State of Texas Education Research pants at Harvara's Program on Laucation Poucy and Governance, the State of Texas Laucation Research

Center at Texas A&M University, the Association for Education Finance and Policy, the NBER Economics

of Education Program LIT Austin's Formation Department and the Stanford Craduate School of Education Center at Texas A&M University, the Association for Education Finance and Policy, the NBEK Economics of Education Program, UT-Austin's Economics Department and the Stanford Graduate School of Education Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinital provided outstanding against the Colin Sullivan Heather Sanone Shalks Lin and Manat Latterinitation and Manat Latterinitation and M of Education Program, UT-Austin's Economics Department and the Stanford Graduate School of Education. Colin Sullivan, Heather Sarsons, Shelby Lin, and Napat Jatusripitak provided outstanding research tion. Colin Sullivan, Heather Sarsons, Shelby Lin, and Napat Jatusripitak provided outstanding R3054 120466 non. Coun Sunvan, Heatner Sarsons, Snelby Lin, and Napat Jatusripitak provided outstanding research assistance. This research was funded by the Institute of Education Sciences under award R305A120466.

Institutional support from Torge A. & M. University, Stanford's Content for Education Policy Analysis and assistance. This research was funded by the Institute of Education Sciences under award R305AI20466.

Institutional support from Texas A&M University, Stanford's Center for Education Policy Analysis, and Harvard's Taubman Center for State and Local Covernment are also gratefully administrated. Institutional support from Texas A&M University, Stanford's Center for Education Policy Analysis, and Harvard's Taubman Center for State and Local Government are also gratefully acknowledged. The data used in this article can be obtained beginning June 2015 through May 2018 from the authors. [Submitted July 2013; accepted Pebruary 2014]
ISSN 0022-166X E-ISSN 1548-8004 © 2015 by the Board of Regents of the University of Wisconsin System

THE JOURNAL OF HUMAN RESOURCES • 50 • 1





Opportunity cost



The arts make kids better at school



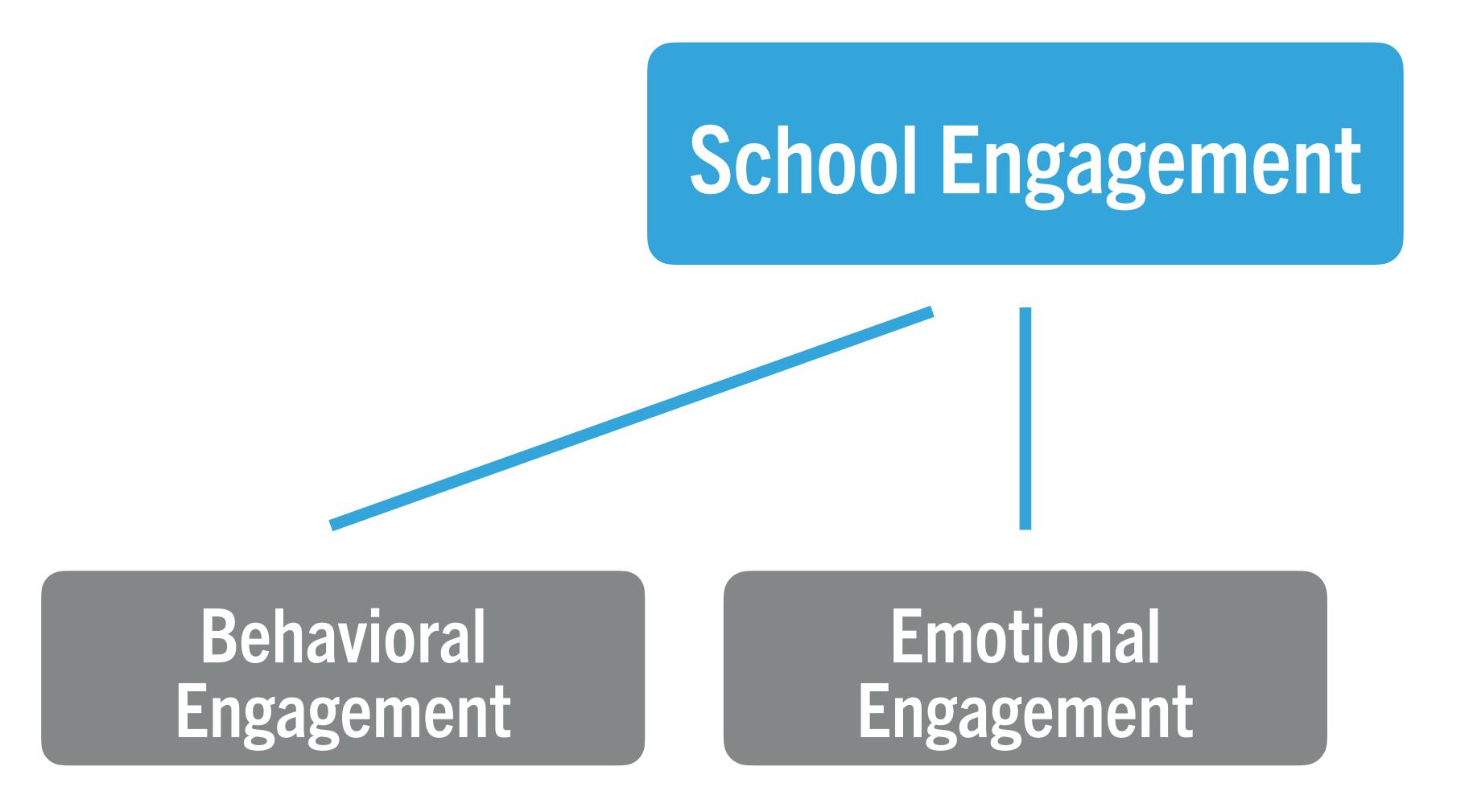
How might the arts make a difference?

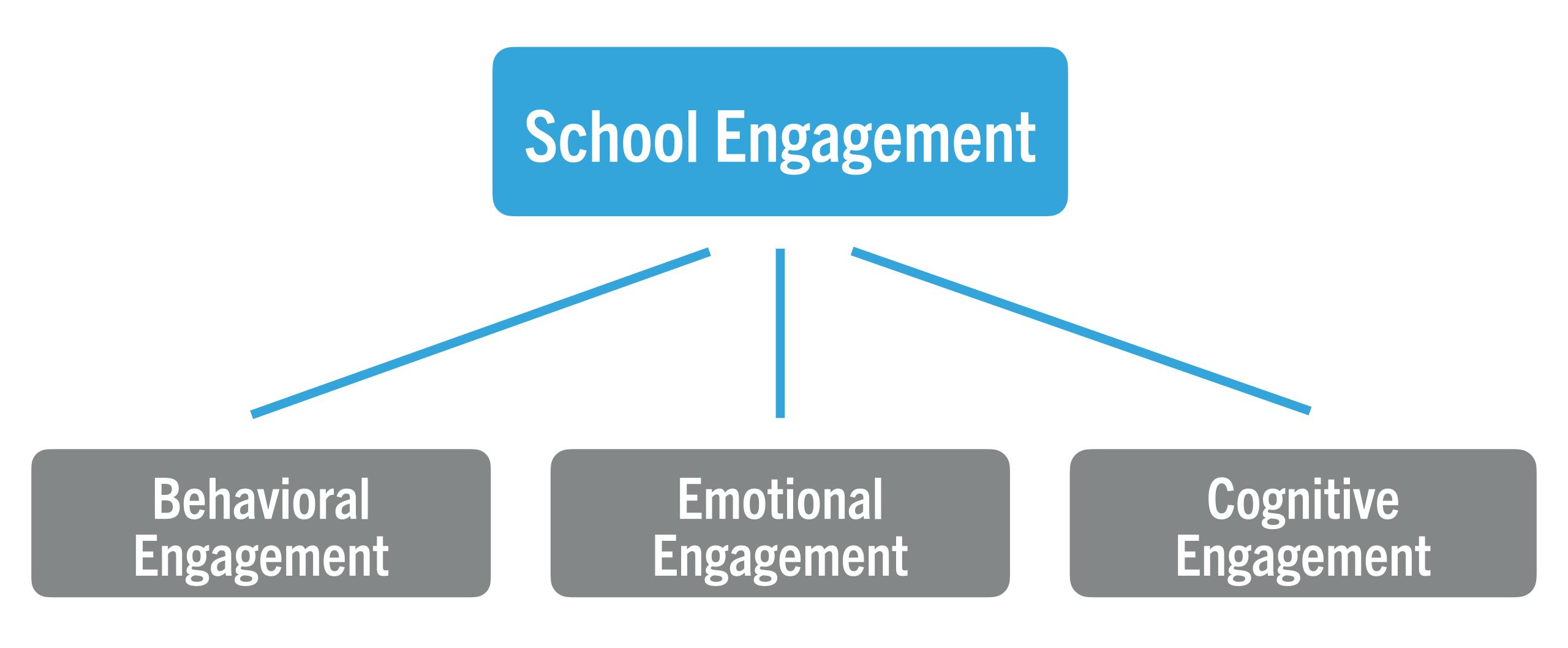


School Engagement

School Engagement

Behavioral Engagement





School Engagement

Academic Outcomes

School Engagement

Academic Outcomes

Better Grades or Test Scores

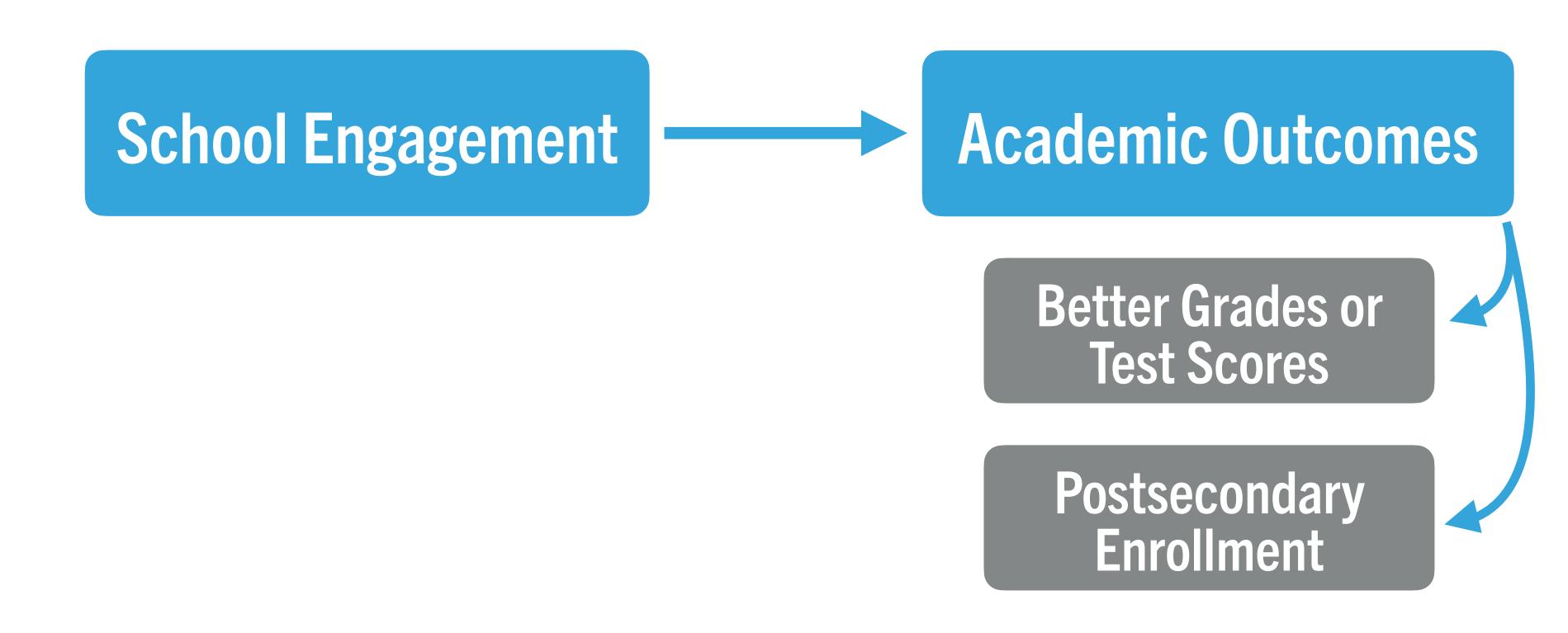


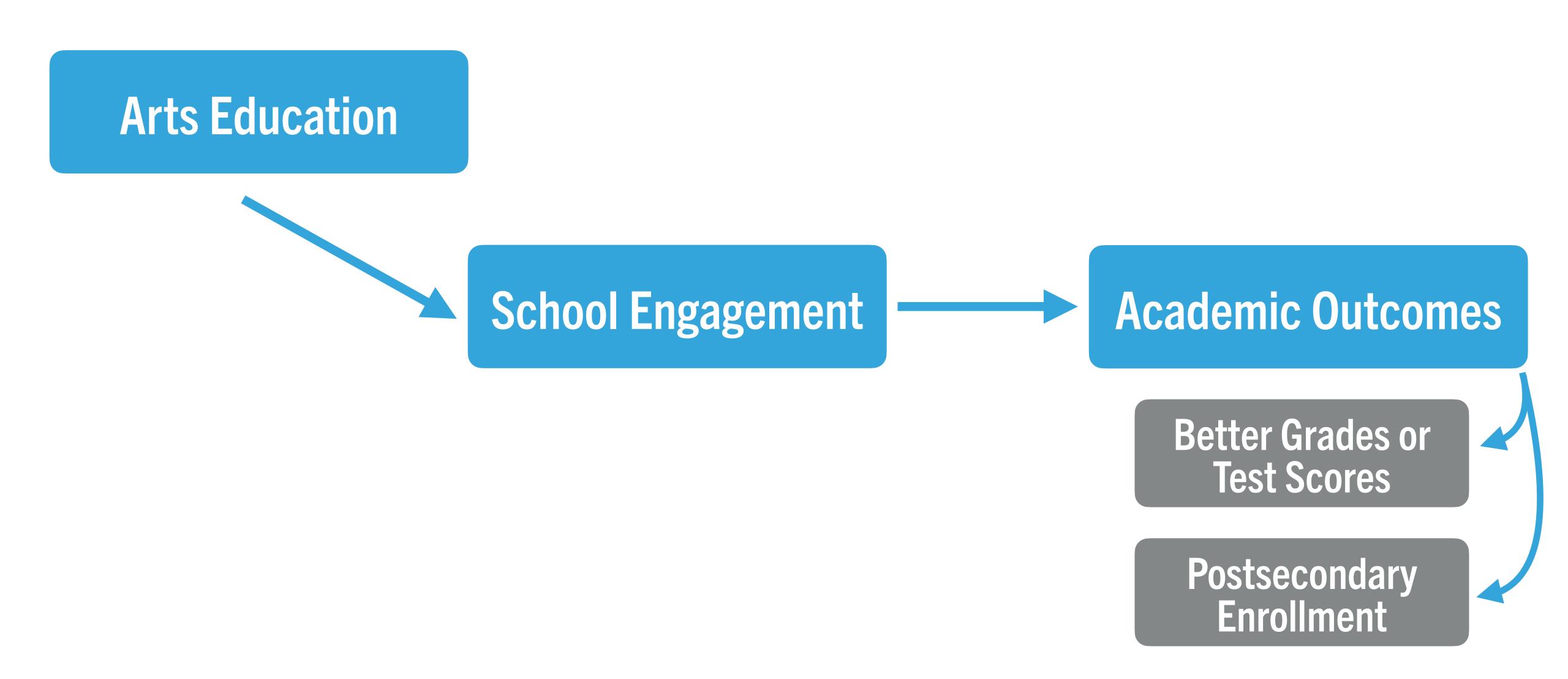


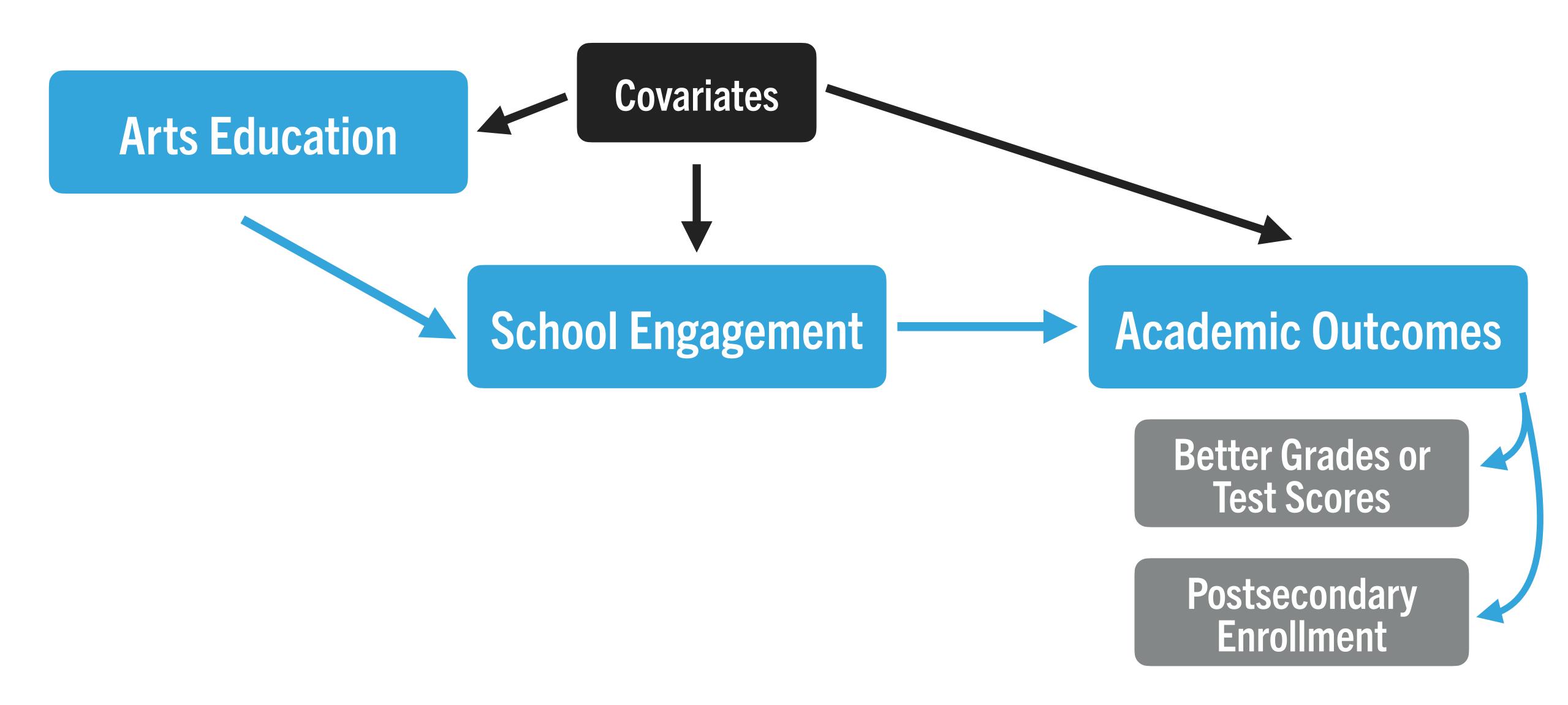
Academic Outcomes

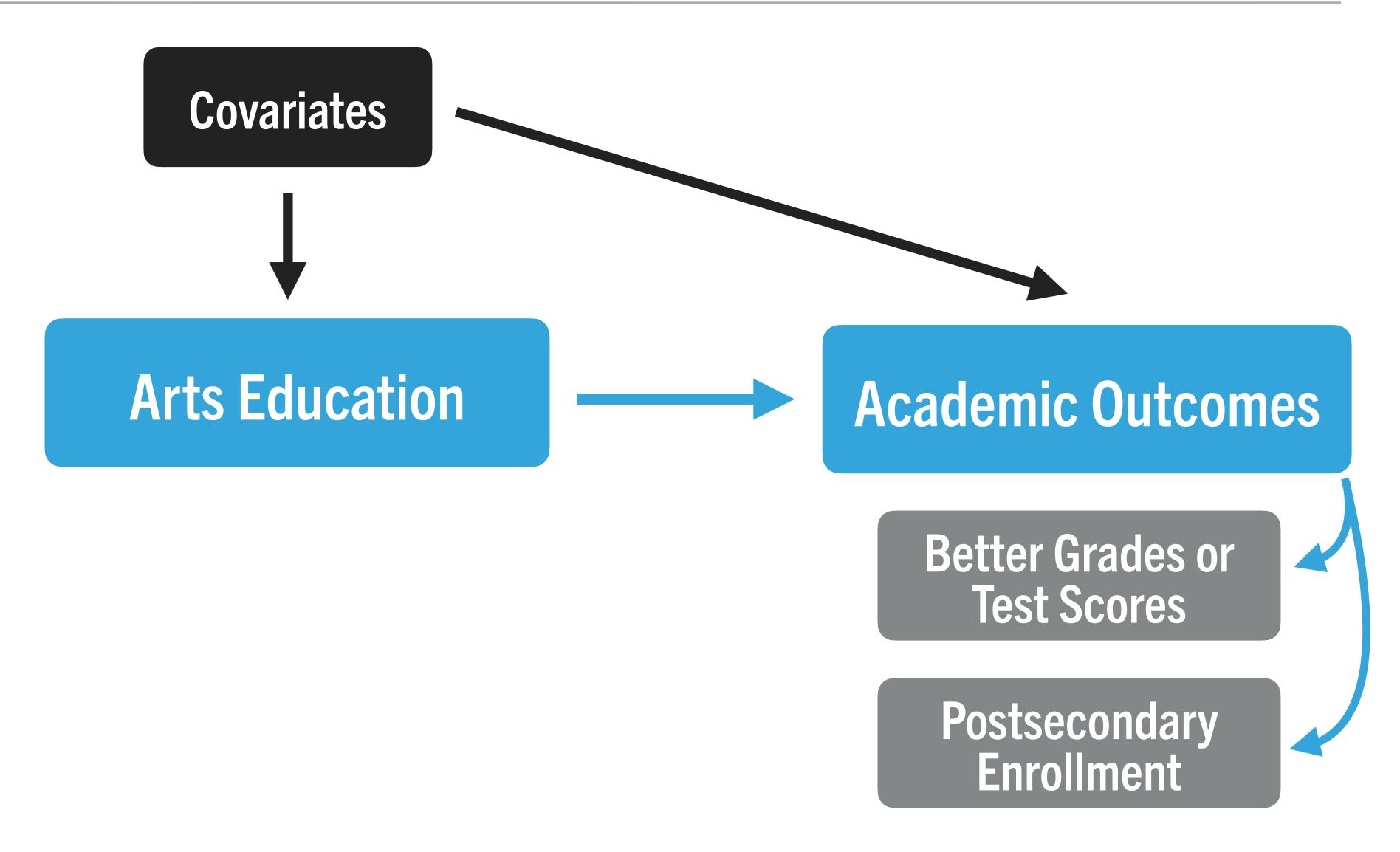
Better Grades or Test Scores

Postsecondary Enrollment











Some Problems with the Evidence Base

Some Problems with the Evidence Base

Unclear Definition

Some Problems with the Evidence Base

Unclear Definition

Varying Curricula

Unclear Definition

No Arts Measures

Varying Curricula

Unclear Definition

No Arts Measures

Varying Curricula

No Fidelity Measures

Unclear Definition

No Arts Measures

No Common Teacher Preparation

Varying Curricula

No Fidelity Measures

Unclear Definition

No Arts Measures

No Common Teacher Preparation

Varying Curricula

No Fidelity Measures

Varying Teacher Quality



If we could find a common arts curriculum administered across many schools with teachers specifically trained to deliver it, we could improve the evidence base.



International Baccalaureate Baccalauréat International Bachillerato Internacional



International Baccalaureate[®]
Baccalauréat International
Bachillerato Internacional



International Baccalaureate Baccalauréat International Bachillerato Internacional

Common Definition



International Baccalaureate[®]
Baccalauréat International
Bachillerato Internacional

Common Definition

Identical Curricula



International Baccalaureate[®]
Baccalauréat International
Bachillerato Internacional

Common Definition

Common Teacher Preparation

Identical Curricula



International Baccalaureate Baccalauréat International Bachillerato Internacional

Common Definition

Arts Measures

Common Teacher Preparation

Identical Curricula



International Baccalaureate Baccalauréat International Bachillerato Internacional

Common Definition

Arts Measures

Common Teacher Preparation

Identical Curricula

Fidelity Assurance



International Baccalaureate[®]
Baccalauréat International
Bachillerato Internacional

Common Definition

Arts Measures

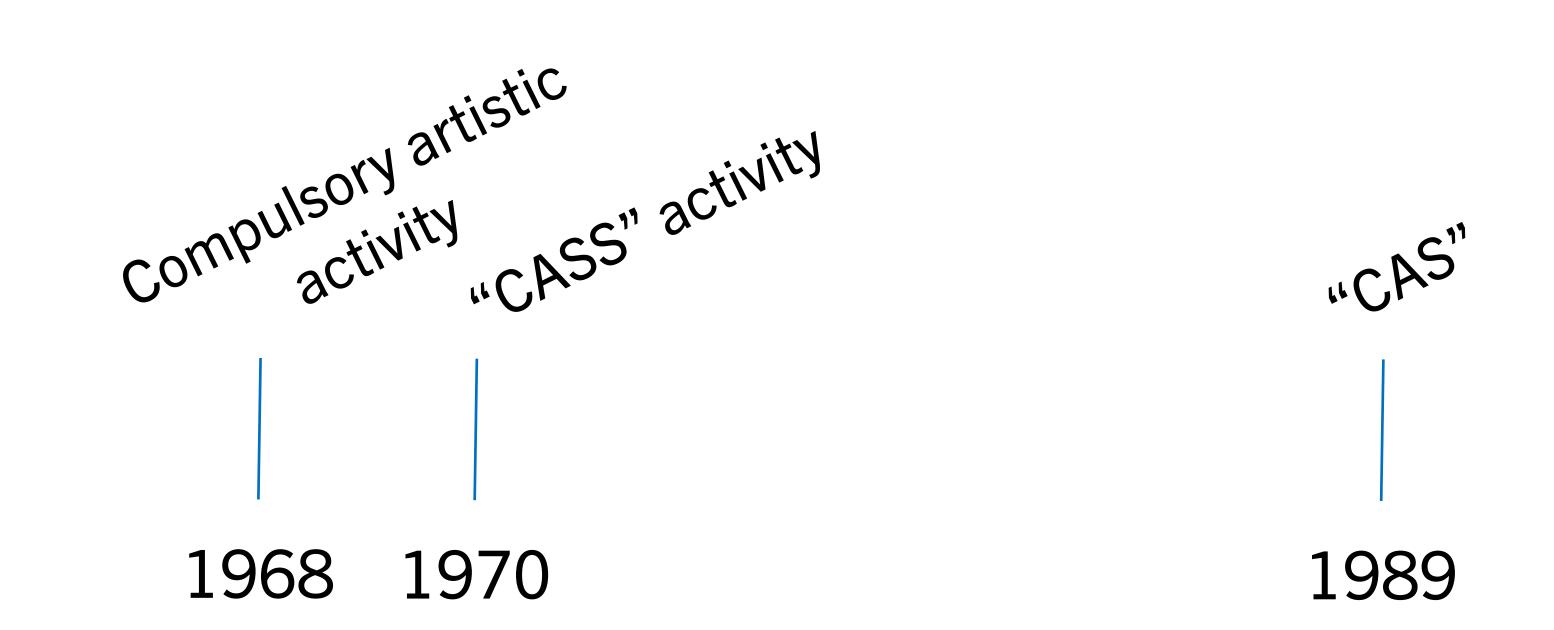
Common Teacher Preparation

Identical Curricula

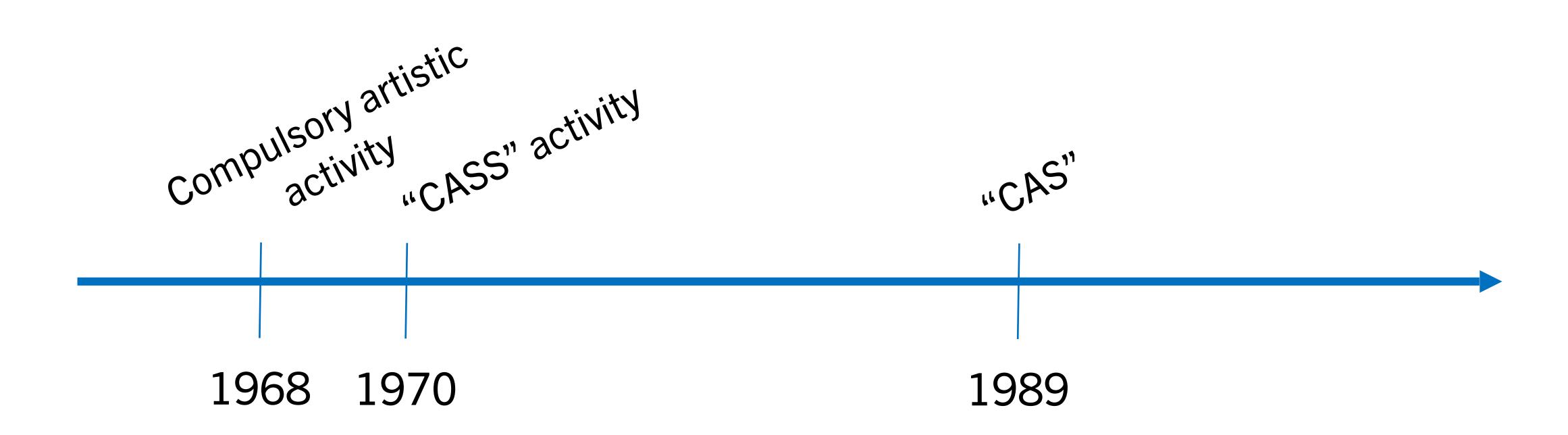
Fidelity Assurance

Varying Teacher Quality

Some form of aesthetic engagement has always been part of the IB ethos.



Some form of aesthetic engagement has always been part of the IB ethos.



Inquirers

Open Minded



Knowledgeable

Caring

Thinkers

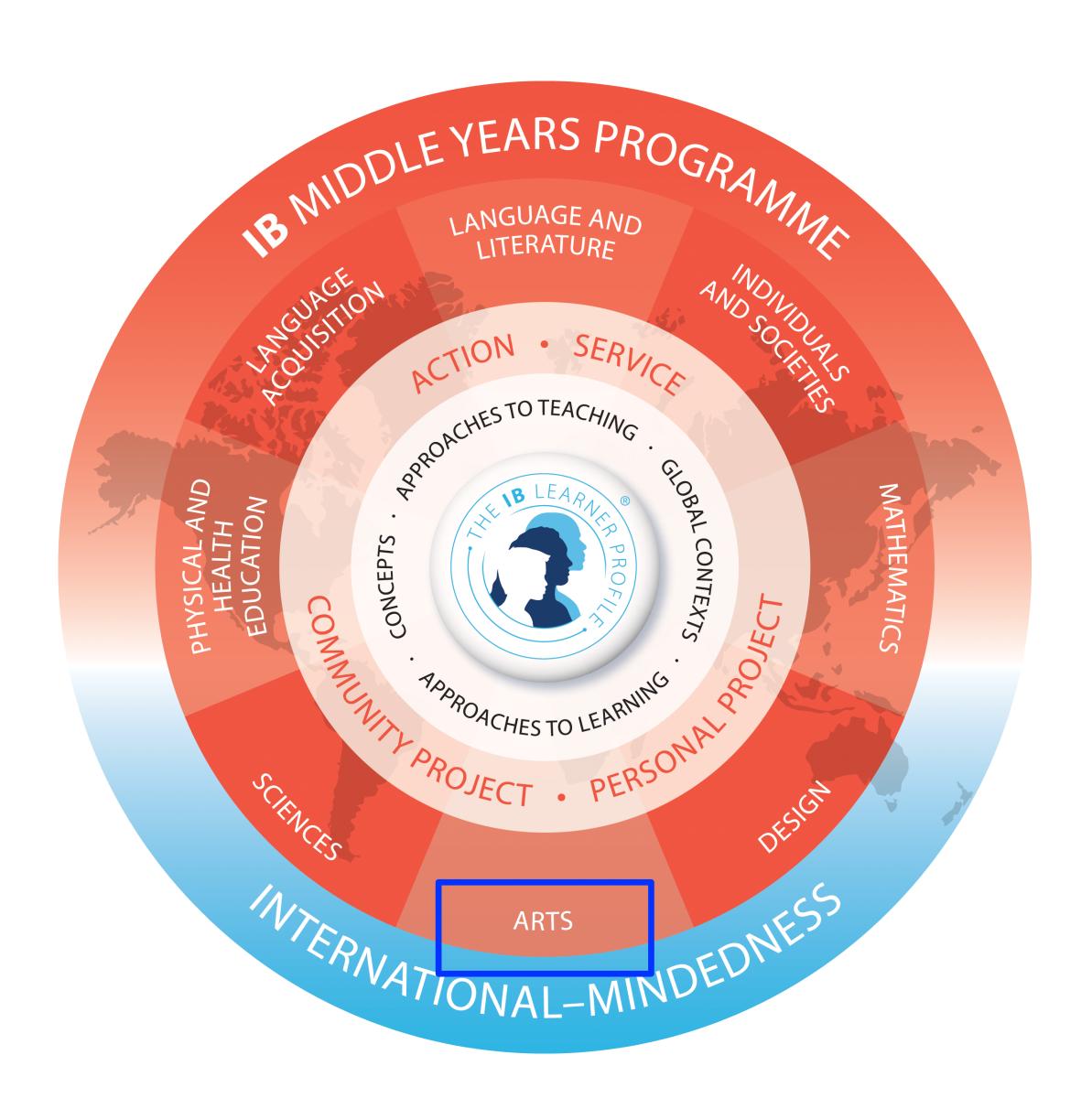
Risk-Takers

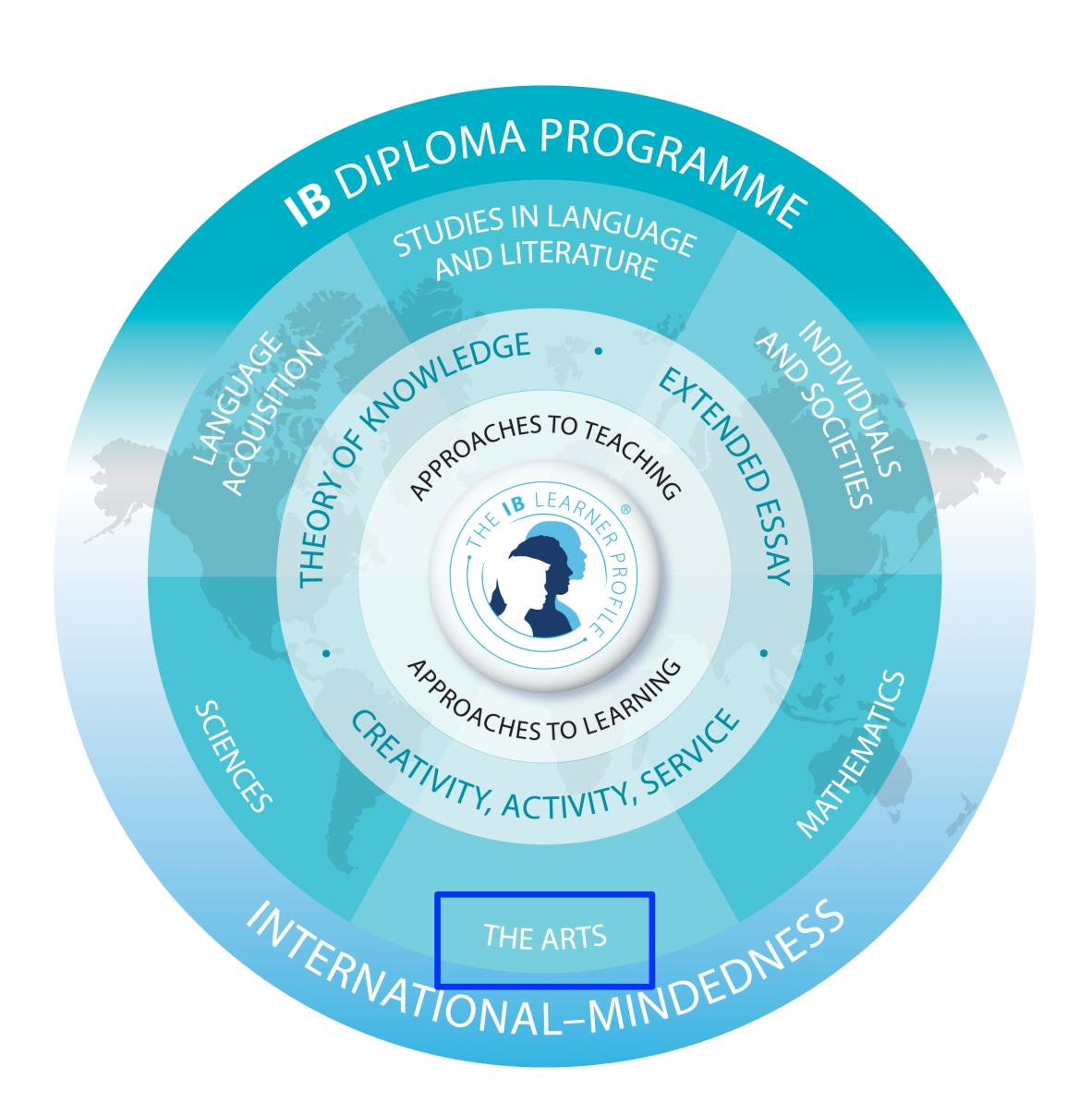
Communicators

Balanced

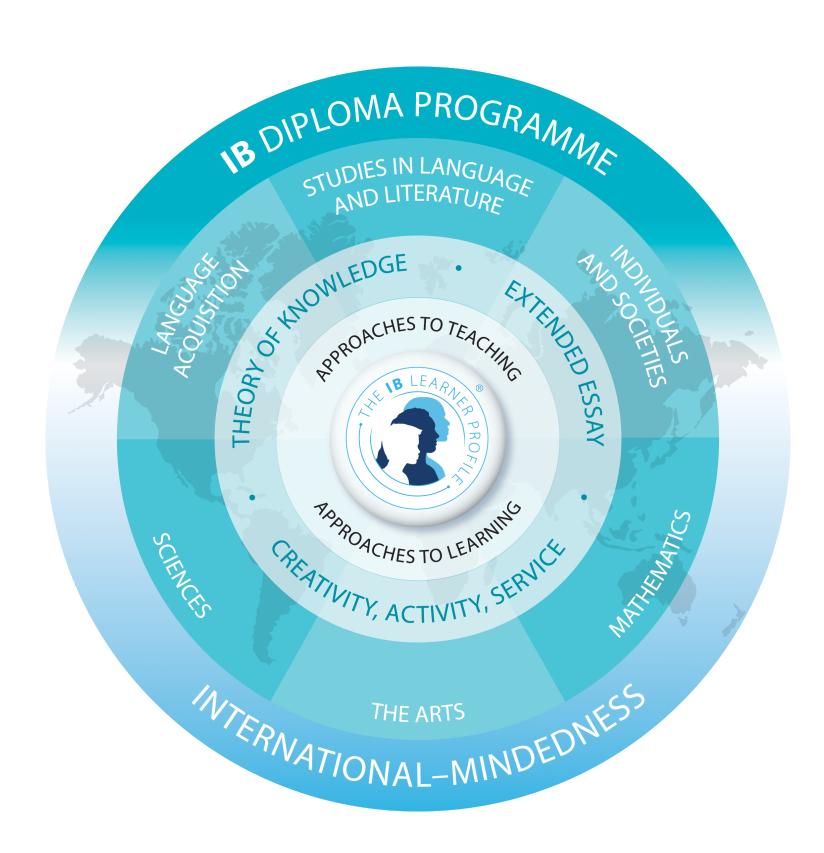
Reflective

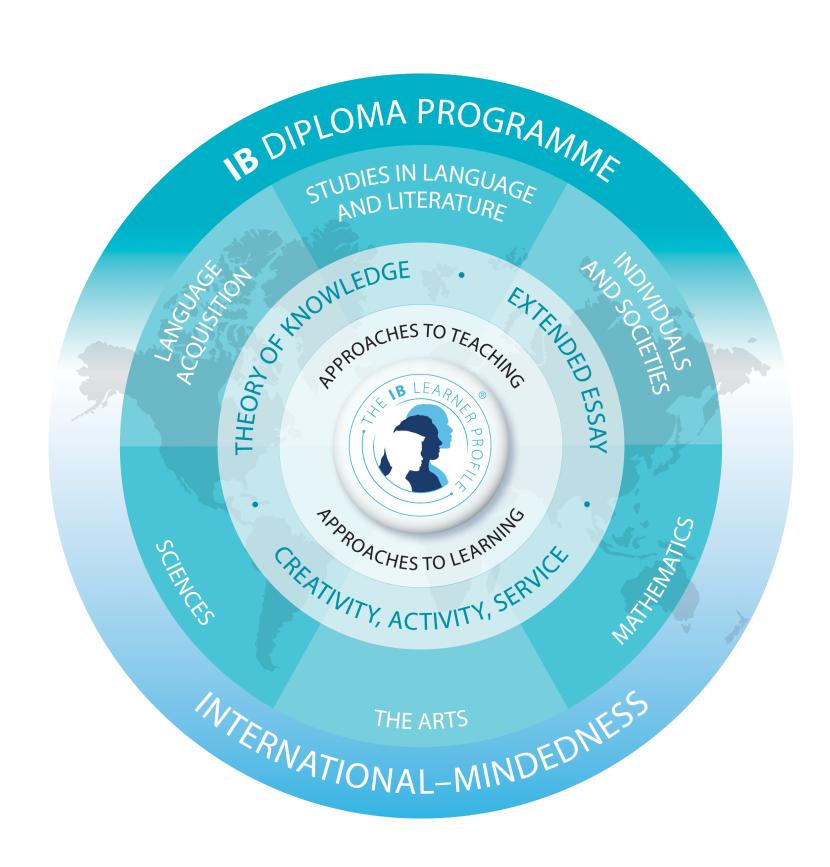
Principled



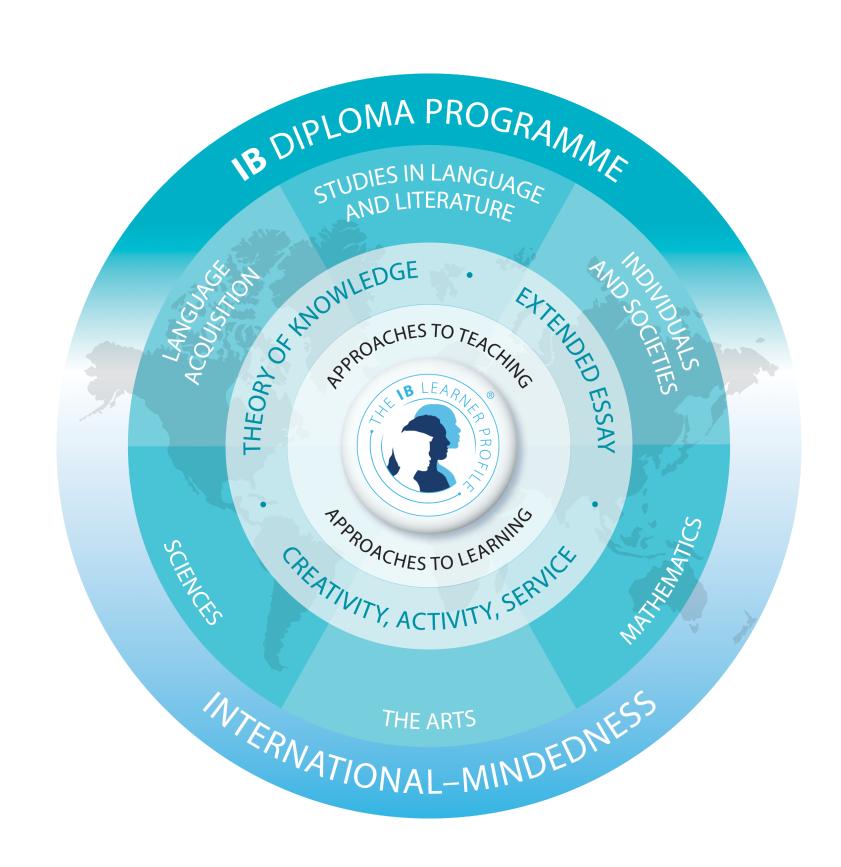








23 Public High Schools in Maryland Authorized to offer the IB Diploma Program across 8 Local School Systems:



23 Public High Schools in Maryland Authorized to offer the IB Diploma Program across 8 Local School Systems:

Anne Arundel County

Baltimore City

Baltimore County

Frederick County

Harford County

Montgomery County

Prince George's County

Washington County



Focus on Music

Focus on Music

Small Sample Sizes

Focus on Music

Small Sample Sizes

Focus on Music

Settings With Low External Validity

Small Sample Sizes

This study...

Focus on Music

Settings With Low External Validity

Small Sample Sizes

This study...

All IB Arts Disciplines

Settings With Low External Validity

Small Sample Sizes

This study...

All IB Arts Disciplines

Settings With Low External Validity

Decent Sample Size

This study...

All IB Arts Disciplines

Settings With Low External Validity

Decent Sample Size

Propensity Score Model

This study...

All IB Arts Disciplines

Authentic School Data

Decent Sample Size

Propensity Score Model

We can exploit the elective nature of the arts in the IB DP to compare outcomes of arts and non-arts students in a more rigorous way than some prior research.

The purpose of this study is to explore the academic achievement of IBDP Arts students in Maryland to understand the relationship between arts study and academics in the IBDP.

Controlling for population differences, do DP arts students outperform DP nonarts students on exams in Math?

Does postsecondary enrollment vary among former IBDP arts and non-arts students?

Do IBDP Arts and Non-Arts Students pursue postsecondary STEM majors at similar or dissimilar rates?

Stratified Propensity Score Model

N = 810

Maryland Students enrolled in full IBDP during 10th and 11th Grade

2015 Graduation Cohort

N = 810

Maryland Students enrolled in full IBDP during 10th and 11th Grade

2015 Graduation Cohort

30% enrolled in an IB Arts course n = 250

N = 810

Maryland Students enrolled in full IBDP during 10th and 11th Grade

2015 Graduation Cohort

30% enrolled in an IB Arts course n = 250

70% did not enroll in an IB Arts course n = 560

K12 Course Enrollment

K12 Course Enrollment

K12 Assessment

K12 Course Enrollment

K12 Assessment

Postsecondary Enrollment

Arts & Non-Arts Sample Characteristics Before and After Propensity Score Stratification

	Sample Proportions (Unmatched)		Standardized Differences	
Characteristic	IB Arts	IB Non-Arts	Unmatched	Matched
Female	0.66	0.61	0.119	0.015
White	0.35	0.39	-0.086	0.006
Black	0.37	0.24	0.292	0.000
Asian	0.13	0.22	-0.225	0.007
Hispanic or Latino	0.10	0.12	-0.048	-0.017
Other Race/Ethnicity or Multiracial	0.06	0.05	0.034	0.000
Free or Reduced Meals (FARMS)	0.28	0.22	0.137	0.004
Special Education Services	0.03	0.03	0.030	0.018
English Language Learner	0.03	0.03	-0.002	0.001
Prior Arts Experience	0.93	0.89	0.125	0.006
Prior HS GPA	3.45	3.55	-0.181	-0.007
School Size	1674.10	1747.41	-0.208	0.006
School FARMS Proportion	0.42	0.35	0.454	0.007
Ν	250	560		

Doubly-robust within-stratum estimation averaged across all strata

Math Outcome (OLS):

 $IBMath_{ij} = \alpha + \tau IBArts_{ij} + \beta_1 Female_{ij} + \beta_2 RaceEthnicity_{ij} + \beta_3 FARMS_{ij} + \beta_4 ELL_{ij} + \beta_5 SpecEd_{ij} + \beta_6 PriorArts_{ij} + \beta_7 PriorGPA_{ij} + \beta_8 SchoolSize_j + \beta_9 SchoolFARMSProp_j + \varepsilon_{ij}$

Postsecondary Outcomes (Logistic Regression):

 $\ln[\frac{P(PostSecondary_{ij})}{1 - P(PostSecondary_{ij})}] = \alpha + \tau IBArts_{ij} + \beta_1 Female_{ij} + \beta_2 RaceEthnicity_{ij} + \beta_3 FARMS_{ij} + \beta_4 ELL_{ij} + \beta_5 SpecEd_{ij} + \beta_6 PriorArts_{ij} + \beta_7 PriorGPA_{ij} + \beta_8 SchoolSize_j + \beta_9 SchoolFARMSProp_j + \varepsilon_{ij}$

Results: IB Math Score

IB Math Scores are Reported on a 1 to 7 Scale, M = 4.80, $SD \approx 1.0$

Doubly Robust Estimator, Average Treatment Effect

	b	SE	P
IB Arts	-0.36	0.092	.001

Results: Postsecondary Enrollment

Doubly Robust Estimator, Average Treatment Effect

	Odds Ratio	SE	P
IB Arts	0.77	0.15	.197

Results: Choice of a STEM Major in Postsecondary School

Doubly Robust Estimator, Average Treatment Effect

	Odds Ratio	SE	p
IB Arts	0.77	0.201	.323

Choice constraints on IB students may channel less academically apt IB students into the IB Arts classes, in contrast to other contexts previously studied.

Students who pursued an IB Arts course enrolled in postsecondary institutions and majored in STEM fields at rates indistinguishable from non-IB Arts students.

School counselors and others should be aware that foregoing arts coursework in favor of other academics may not result in better postsecondary transitions.

Limitations of the Study

- · Single Cohort
- Strong ignorability assumption may not hold
- Self-selection into IB may limit external validity to other populations



Kenneth Elpus David Miller elpus@umd.edu dmille20@umd.edu

