SAT Scores of Students Who Study the Arts: What We Can and Cannot Conclude about the Association

KATHRYN VAUGHN and ELLEN WINNER

> Music actually makes our kids smarter . . . . The College Board last year documented a 100-point gap in SAT scores between students who had music instruction during their early elementary school years and students who did not. The longer students study music, the greater the gap in scores.
>
> —From testimony by the head of a cable TV station along with Wynton Marsalis presented to the U.S. House of Representatives Education Caucus, July, 1999.

Perhaps the finding most often mentioned in the popular press about the relationship between arts and academic achievement is that students who study the arts in high school have higher SAT scores than those who do not study the arts. This kind of claim is exemplified by the quotation above, presented as testimony to the U.S. House of Representatives. The relationship between arts courses and SAT scores has been documented by the College Board since 1987 and is based on a very large sample—all students taking the SAT who voluntarily responded to the Student Descriptive Questionnaire (SDQ) as part of the registration process. Between 94%-95% of students taking the SAT fill out at least one item on the questionnaire. The College Board considers the data to be "nearly a complete and accurate description of the tested population."[1] The analysis we present here is based on the actual mean scores of students responding to a given question. We have analyzed the College Board data to answer a variety of comparative questions.

The Relationship between SAT Scores and Each Added Year of Arts Study

Item 1 of the SDQ asks, "Indicate the total number of years of high school courses (in grades 9-12) [in Arts and Music] you have taken or plan to take

Journal of Aesthetic Education, Vol. 34, Nos. 3-4, Fall/Winter 2000

in each of the subjects listed below. If you have not taken any course in a subject and do not plan to take any in high school, fill in the oval 'None' column. If you repeat a course, count it only once." The questionnaire lists examples of what is meant by Arts and Music: art, music, art history, dance, and theater. In the following analyses, we compared scores of students who indicated having taken zero, one, two, three, and four years of such arts courses.

We analyzed 12 available years of SAT data, from 1987-1998, examining amount of arts course experience as the between-subjects factor.[2] Because 1996-98 scores were recentered, they were higher than earlier years. Prior to analysis, we converted these scores back to the original scale so that scores were comparable across years. We repeated this analysis three times, once for composite scores and then again for the verbal and math scores separately in order to determine which of these two types of scores is more highly associated with arts study.

We treated each year's score as an independent observation. There were 12 observations for zero, one, two, and three years of arts courses taken, and 8 observations for four and over four years of arts courses taken.[3] These same levels were used for analyses on verbal and math SAT scores, reported later.

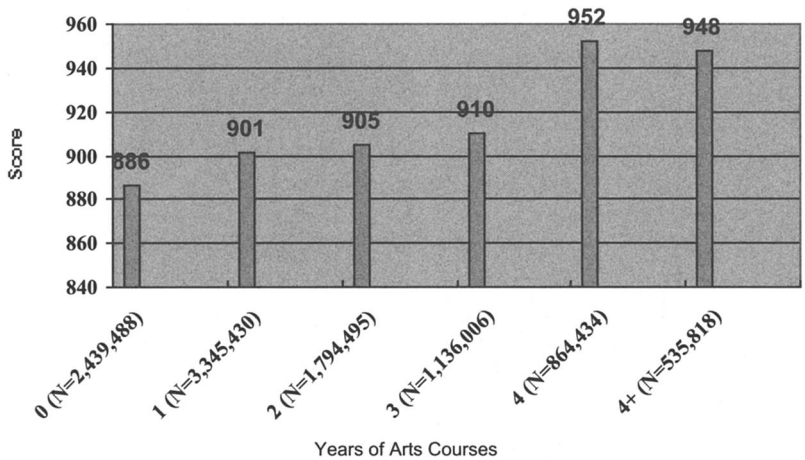
Composite Scores

Figure 1 shows mean composite SAT scores achieved by students with different levels of arts coursework. As can be seen, scores appear to increase gradually with each level of arts from 0-3 years, and then to jump sharply at four years. A one-way ANOVA, with levels of arts experience as the between-subjects factor, revealed a significant effect of level of arts experience, $F(5,60)=93.977$, $MSE=83.791$, $p<.0001$. A contrast analysis revealed scores of students with zero, one, two, and three years of arts experience to be significantly lower than scores of students with four and over four years experience, $F(1,60)=426.544$, $MSE= 83.791$, $p<.0001$. This contrast alone accounted for 91% of the variability across the different levels of arts education.

A trend analysis was then performed on scores of students with zero to three years of arts courses. The error term from the original ANOVA was used for this and other tests for trend components. A significant linear component was revealed, $F\ (1,60)=43.009$, $MSE=83.791$, $p<.0001$, reflecting the fact that scores increase linearly from 0-3 years of arts experience.

Thus, composite SAT scores increase linearly from zero to three years of arts experience, and then rise sharply at four years. While scores of students with over four years of arts are slightly lower than those with four years, a post-hoc *t*-test showed that the mean scores of students with four and over four years of arts were not reliably different, $p>.05$.

Figure 1
Composite SAT Score as a Function of High School Arts Courses



Verbal Scores

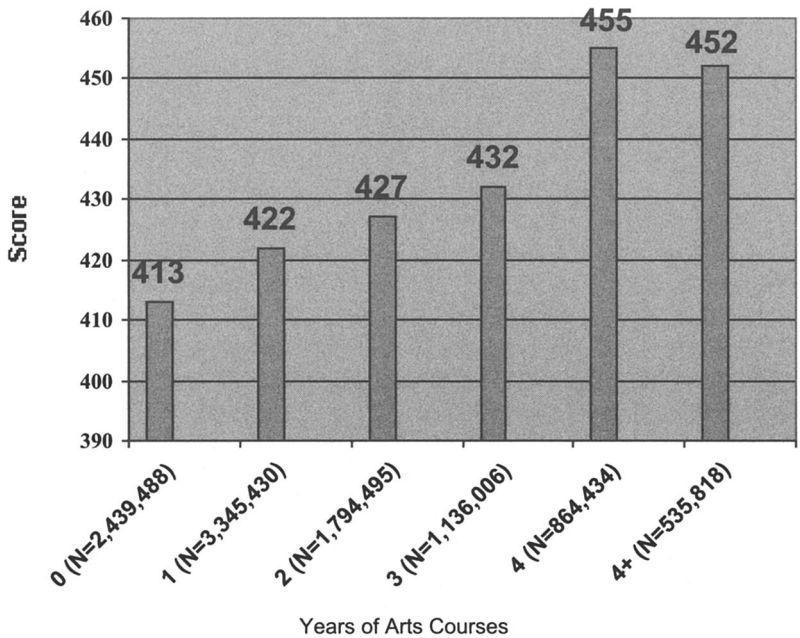
Figure 2 shows the relationship between verbal SAT scores and years of high school arts courses. A one-way ANOVA revealed a significant effect of level of arts experience, $F(5,60)=167.960$, $MSE=18.02$, $p<.0001$. A contrast analysis revealed scores of students with zero, one, two, and three years of arts experience to be significantly lower than scores of students with four and over four years experience, $F(1,60)=715.824$, $MSE=18.021$, $p<.0001$. This contrast alone accounted for 85% of the variability across the different levels of arts education.

A trend analysis performed only on scores of students with zero to three years of arts revealed a significant linear component, $F(1,60)=132.844$, $MSE=18.021$, $p<.0001$. No other component accounted for any significant amount of variance. Thus, as with composite scores, verbal SAT scores increase linearly from zero to three years of arts experience, and then rise sharply at four years. And as with composite scores, while scores of students with over four years of arts are slightly lower than those with four years, a post-hoc *t*-test showed that the scores of these two groups of students were not reliably different, $p>.05$.

Math Scores

Figure 3 shows mean math SAT scores achieved by students with different levels of arts coursework. As with the composite and verbal scores, there is again a sharp discontinuity between three and four years. A one-way ANOVA revealed a significant effect of level of arts experience, $F(5,60)=30.175$, $MSE=39.438$, $p<.0001$. A contrast analysis revealed scores of students with zero, one, two, and three years of arts experience to be significantly

Figure 2
Verbal SAT Score as a Function of High School Arts Courses



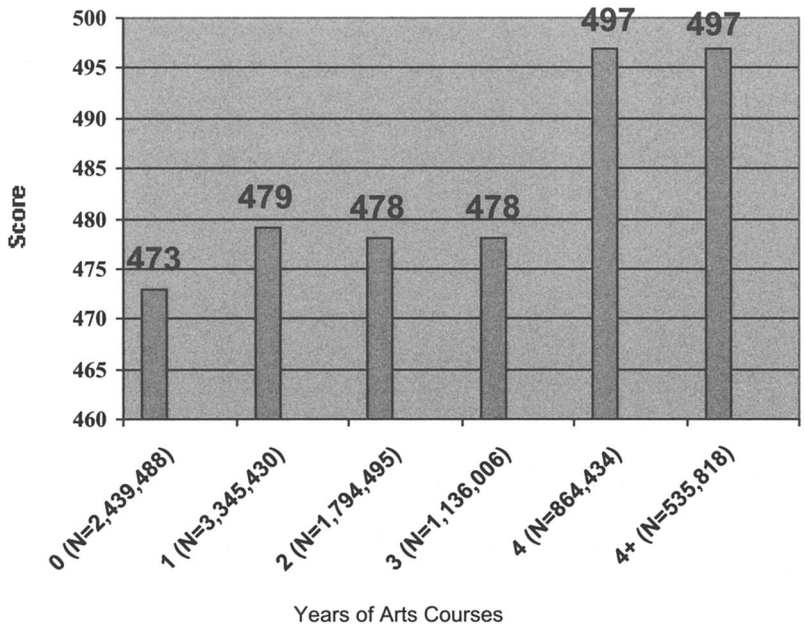
lower than scores of students with four and over four years experience, $F(1,60)=144.434$, $MSE=39.438$, $p<.0001$. This contrast alone accounted for 96% of the variability across the different levels of arts education.

A trend analysis performed only on scores of students with zero to three years of arts revealed a nearly significant linear component, $F(1,60) = 3.126$, $MSE=39.438$, $p=.08$. No other component accounted for any consequential amount of variance. As with composite scores, while scores of students with over four years of arts were slightly lower than those with four years, a post-hoc *t*-test showed that these means were not reliably different, $p>.05$

Which Test Is More Strongly Associated with Studying the Arts: The Verbal or the Math SAT?

We next combined scores meta-analytically, once for verbal scores and once for math scores. Each meta-analysis was based on ten years of SAT data, 1988-1998 (with 1993 missing).[4] Effect sizes were computed for each year comparing students with 4 vs. 0 years of arts. The mean weighted effect size for the verbal scores was $r=.19$ (Stouffer's $Z=363.92$, [5] $p<.0001$), with a 95%

Figure 3
Math SAT Score as a Function of High School Arts Courses

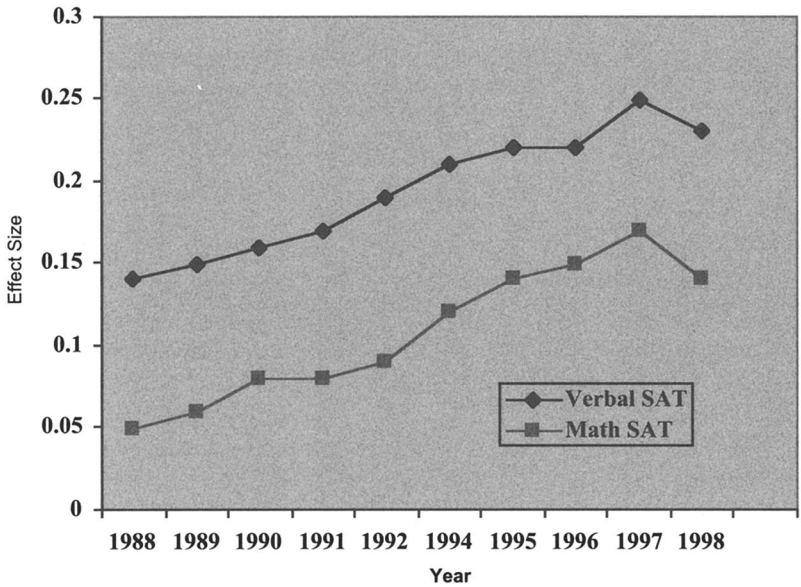


confidence interval of r=.17 to r=.22. We can thus conclude that there is a significant relationship between higher verbal scores and taking 4 years of arts. However, the 10 effect sizes were significantly heterogeneous χ^2 =4963, df=9, p <. 001.[6] As reported in the article by Winner and Cooper in this issue, a linear contrast test showed that effect sizes rose over the years.

A comparable meta-analysis performed on 10 years of math SAT data yielded a mean weighted and unweighted effect size of r=.11, with a 95% confidence interval of r=.08 to r=.13 (Stouffer's Z=206.70, p<.001). These math effect sizes are considerably smaller than the verbal effect sizes. Again, effect sizes were significantly heterogeneous χ^2=5,396, df=9, p<.001. And again, as reported in the article by Winner and Cooper in this issue, a linear contrast test revealed that effect sizes rose over time.

Figure 4 displays the rise in effect sizes over ten years for both verbal and math scores. Note that the fact that effect sizes grew over the years is unrelated to recentering of scores, since our effect sizes were computed on scores adjusted back to nonrecentered scores. Figure 4 also clearly shows how the effect sizes for math scores are consistently smaller than those for verbal scores.

Figure 4
Effect Sizes Comparing Zero to Four Years of Arts Courses for Ten Years of Verbal and Math SAT Data



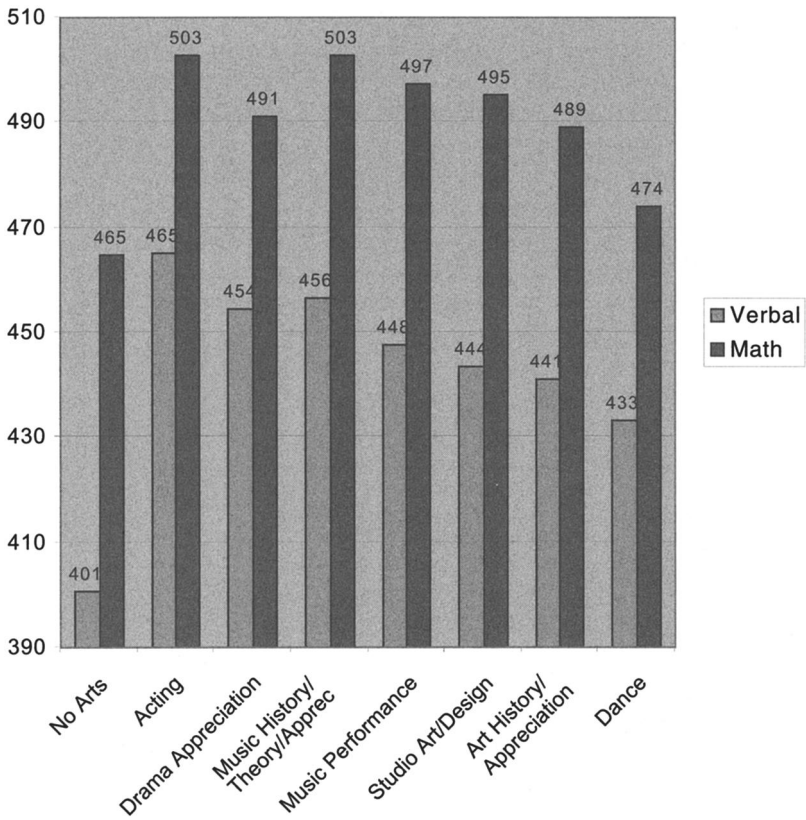
Does It Matter which Art Form?

A later question on the SDQ asks, "Please provide information about the content of some of the high school courses that you have taken or plan to take, and related activities. You may mark more than one in each subject area." Students were then given the following choices: No course work in the arts; acting or the production of a play; drama or theater appreciation; studio art and design; art history or art appreciation; dance; music history, theory, or appreciation; music, instrumental or vocal performance; photography or filmmaking. We examined 10 years of data relating SAT scores to each of the above types of arts courses (excluding photography/filmmaking).[6]

We first compared verbal scores (averaged over 10 years) across art form. As can be seen in Figure 5, among students taking some form of art, the highest scores were obtained by those who took one or more courses in acting/play production (M=465), and the lowest scores were obtained by students who took one or more courses in dance (M=433). Note that the lowest verbal scores were obtained by students who took no arts at all (M=401).

We next compared math scores (averaged over 10 years) across art form. As can also been seen in Figure 5, among students taking some form of art,

Figure 5
Ten Year Average of Mean SAT Scores Associated with Specific Arts Course Participation (Unrecentered Scores)



the highest scores were obtained by students taking acting/play production courses and music history/theory/appreciation (M=503), and the lowest scores were once again obtained by students taking dance (M=474). Note again that the lowest math scores were obtained by students who took no arts at all (M=465).[7]

t-test Comparisons. *t*-tests, performed separately for verbal and math scores, compared no-arts scores to scores for each type of arts course. These tests are summarized in Tables 1 (verbal) and 2 (math). All of the statistical comparisons proved significant. Thus, we can conclude that the verbal and math SAT scores of students taking any form of art, irrespective of number of years, are significantly higher than for students who take no art.

Effect Size Comparisons. We next computed an effect size over the ten years of data comparing the scores of students taking no arts vs. the scores of

Table 1: T-tests Comparing Mean Verbal SAT Scores Over 10 Years: No Arts vs. Specific Arts Courses (Arts Courses Listed in Descending Order of Means)

Art Form	Mean	SD*	df	*t*	one-tailed *p*
No Arts	400.7	2.95			
Acting	465.2	3.01	18	48.44	$p<.0001$
Music History/Theory/Apprec	456.4	4.27	18	34.36	$p<.0001$
Drama Appreciation	454.3	2.75	18	42.05	$p<.0001$
Music Performance	447.6	3.80	18	31.28	$p<.0001$
Studio Art	443.6	2.63	18	34.87	$p<.0001$
Art History/Appreciation	441.0	3.43	18	28.16	$p<.0001$
Dance	433.1	2.84	18	25.56	$p<.0001$

Note: Standard deviations are calculated using each year as an instance for the particular art form.

Table 2: T-tests Comparing Mean Math SAT Scores over 10 Years: No Arts vs. Specific Arts Courses (Arts Courses Listed in Descending Order of Means)

Art Form	Mean	SD	df	*t*	one-tailed *p*
No Arts	464.9	2.80			
Music History/Theory/Apprec	503.7	8.12	18	14.28	p<.0001
Acting	502.9	6.29	18	17.45	p<.0001
Music Performance	497.2	7.31	18	13.04	p<.0001
Studio Art	495.4	7.58	18	11.92	p<.0001
Drana Appreciation	491.0	8.06	18	9.56	p<.0001
Art History/Appreciation	489.2	5.67	18	12.15	p<.0001
Dance	477.5	7.67	18	4.87	p<.0002

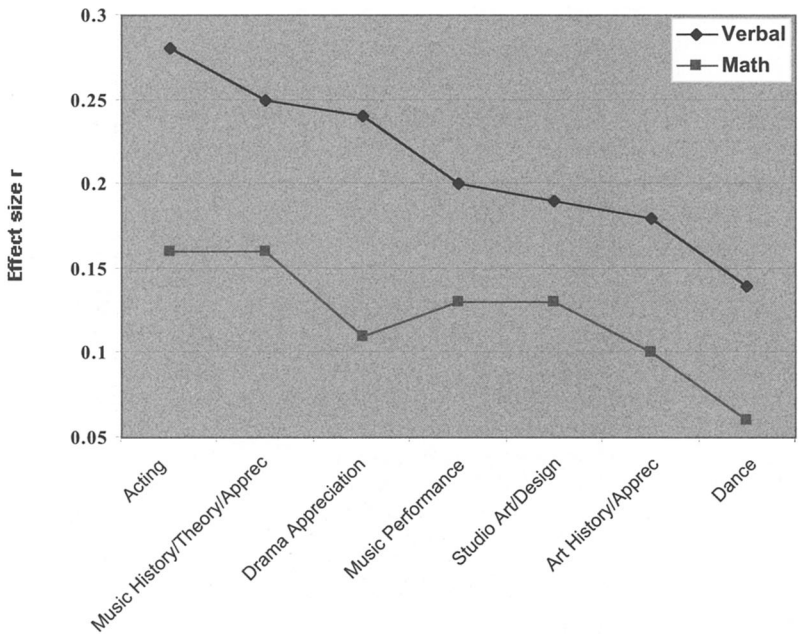
Note: Standard deviations are calculated using each year as an instance for the particular art form.

students taking a specific art form. We computed effect sizes separately for each art form because the groups were not independent. That is, students who took one form of art may well also have been represented among those taking another form.[8]

We conducted independent meta-analyses for each art form and compared the weighted average effect sizes yielded by each one. Thus, we performed a total of 14 small meta-analyses (seven art forms by two outcomes). These analyses enabled us to determine whether the relationship between math scores and the experience of taking a music course was stronger than, for example, the relationship between math scores and the experience of taking a visual arts course.

Figure 6 shows the mean weighted effect size for verbal and math SAT and participation in each art form, shown in relative descending order for

Figure 6
Effect Sizes for Mean SAT Scores By Arts Course Participation 1988-98



verbal scores. Effect sizes for verbal scores ranged from a high of r=.28 (acting) to a low of r=.14 (dance). Effect sizes for math scores ranged from r=.16 (for acting) to r=.06 (dance). Although the math effect sizes ranged less widely and were lower overall, a similar pattern was found for both verbal and math scores: the strongest relationship between score and type of arts course obtained when the course was in acting, the lowest obtained when the course was in dance.

Tables 3 and 4 present further information about the strength of the mean effect size found for each art form. In these tables, we first list the effect size, followed by the 95% confidence interval (which indicates the interval within which the average effect size of another sample of 10 similar studies would be likely to be found). The next column lists the Stouffer's Zs, all of which were associated with p levels of <.001, indicating that these results are generalizable to another sample of students who might have been sampled for these studies. The next column lists the t-tests of the mean Zr, all of which are significant at p<.0001. This allows us to generalize the results to new studies. The file drawer column indicates the number of new studies averaging null results (i.e., mean probability levels of .50) that would need to be unearthed in order to bring the probability level of our Stouffer's Z down to p=.05. These tables clearly show a range of effect sizes

Table 3: Mean Weighted Effect Size for Verbal SAT Across Art Form, Listed in Descending Order

Art Form	Effect Size *r*	95% CI	Stouffer's Z	*t*-test of mean *Zr*	File Drawer Effect	Robustness
Acting	.28	.27 to.30	571.81*	34.98*	1,208,277	11.1
Music History/Theory/ Apprec.	.25	.24 to.27	463.30*	23.77*	793,231	9.3
Drama Appreciation	.24	.23 to.26	424.90*	34.34*	667,164	10.9
Music Performance	.19	.18 to.21	465.07*	23.60*	799,301	7.5
Studio Art/Design	.19	.18 to.21	387.68*	23.89*	555,415	7.6
Art History/Appreciation	.18	.17 to.20	367.43*	22.60*	498,905	7.2
Dance	.14	.13 to.16	265.51*	17.02*	273,524	5.5

* Significant at $p<.0001$.

Table 4: Mean Weighted Effect Size for Math SAT Across Art Form, Listed in Descending Order

Art Form	Effect Size *r*	95% CI	Stouffer's Z	*t*-test of mean *Zr*	File Drawer Effect	Robustness
Acting	.16	.14 to .18	323.63*	15.72*	387,033	4.8
Music History/Theory/ Apprec.	.16	.13 to .18	299.01*	12.33*	330,396	3.9
Music Performance	.13	.11 to .15	306.51*	17.02*	347,165	3.8
Studio Art/Design	.13	.11 to .15	256.38*	10.67*	242,892	3.4
Drama Appreciation	.11	.09 to .13	199.39*	9.43*	146,911	3.4
Art History/Appreciation	.10	.09 to .13	206.39*	12.15*	157,409	3.8
Dance	.06	.03 to .08	108.47*	4.45*	43,472	1.4

* Significant at $p<.0001$.

(depending on type of art course taken), from $r=.10$ (the relationship between art history appreciation and math score) to $r=.28$ (the relationship between acting/play production and verbal score).

Conclusion

Our analyses demonstrate that students who take any kind of art course in high school have higher SAT scores (both math and verbal) than students who take no art course at all. Moreover, those who take four years of arts courses have higher scores than those who take some art but less than four years' worth.

However, we cannot conclude from these findings that taking arts courses will result in higher SAT scores because the data are purely correlational and allow no causal inference. The link between SAT scores and arts

courses has many possible explanations besides the possibility that exposure to the arts by itself leads to the kind of cognitive growth that would be reflected in higher SAT scores. Alternative explanations include the possibility that students who choose to study the arts are high achievers to begin with. It is possible that high-achieving students come from families that value both academic achievement and the arts; hence such students not only work hard on academic subjects but also choose to study the arts. It is possible that high academic achievers choose to take arts courses because they believe these courses will increase their chances of admission to a selective college. And it is possible that students who study the arts attend schools that are strong in both arts and academics.

Although the link between SAT scores and study of the arts is positive, an even stronger link exists between SAT scores and study of *academic* subjects. Elliot Eisner has compared SAT scores of students taking 1, 2, 3, 4, 4+ years of English, math, science, history and social science, foreign language, and arts and music.[9] What he finds is striking. In every case, students who take at least four years of a given academic subject score considerably higher on their SATs (both math and verbal) than do students who take only one year of this subject. Moreover, the difference between scores of students with one vs. four or more years of the subject is far smaller in the case of the arts and music than it is in the case of each of the other academic areas examined. These data suggest that students who focus on particular areas are high-achieving students. Focusing on the arts is one indication of high academic achievement. But focusing on academic areas is an even stronger indication of high academic achievement.

Here are a few more examples to make this point. In 1998, the average (nonrecentered) SAT math score obtained by students who took 4 years of arts was M=512. For students who took 4 years of precalculus, the math score was 59 points higher, M=571. The average SAT verbal score obtained by students who took 4 years of arts was M=461. For students who took 4 years of foreign and classical languages, verbal scores were 20 points higher, M= 481.

Thus, while there is a link between studying the arts and SAT scores, this link is only correlational. We do not as yet understand what underlies this correlation. In addition, the link between SAT scores and electing to study academic subjects in a focussed manner is a stronger one. Policymakers should refrain from citing the link between arts and SAT scores both because their claims imply that a causal relationship exists when nothing causal has been established, and also because the link between arts courses and SAT scores is far weaker than the link between academic courses and SAT scores.

Nonetheless, it is striking that students who study the arts are consistently higher academic achievers than students who do not study the arts. It

is worth noting that the best independent schools in the United States have always retained an important place for the arts. It is only in our financially strapped public schools that tight budgets have led to arguments that the arts can be cut because they are not essential. Perhaps our independent schools have implicitly recognized that the arts are a valuable arena of learning, and that the kinds of discipline as well as self-expression gained from the arts are important in their own rights, irrespective of whether these lead to higher SAT scores.[10]

NOTES

1. *College-Bound Seniors Profile of SAT and Achievement Test Takers* (College Board, 1993), p. 111.
2. We were able to analyze 12 years of data for this analysis because we did not need standard deviations of scores. The meta-analyses reported in Winner and Cooper (this issue) required standard deviations, and these were available for only 10 years of data.
3. From 1994-1995, scores for students with 4 and over 4 years of arts courses were combined into one category of 4 or more years. We did not analyze the combined category of "4 or over 4 years" since this category was not used in the eight other years.
4. We were able to use just 10 years of data for our meta-analyses because we had standard deviations only for 1988-1992 and 1994-1998.
5. To obtain our Stouffer's *Z*, we summed the one-tailed p-levels for the studies and then divided the sum by the square root of the number of studies, as recommended by Robert Rosenthal, *Meta-analytic Procedures for Social Research* (Newbury Park, Calif.: Sage Publications, 1991).
6. We excluded photography because we focus here only on the four traditional art forms: music, visual art, drama, and dance.
7. The scores for students who reported taking no arts are slightly different from the scores listed in Figure 3 for students taking no arts. This is due to the fact that Figure 3 is based on responses to the question asking how many years of arts courses had been taken (0, 1, 2, 3, 4, 4+) (Item 1 of the Student Descriptive Questionnaire).
8. We first used the calculation for Hedges *g* because the variances for each group in each of the ten years (excluding 1993) were available. This measure is the difference between the means of the control and experimental groups divided by the pooled standard deviation (see Harris Cooper and Larry Hedges, eds., *The Handbook of Research Synthesis* [New York: Russell Sage, 1994], p. 233.) We then converted the *g* to Pearson's $r = \frac{g}{\sqrt{g^2 + 4(\frac{.5(n_1 + n_2)}{n_{harmonic}})}}$ (Robert Rosenthal, 1999). All subsequent calculations were performed using the Fisher's *Zr* transformation effect size *r* to account for the non-normal distribution of the *r*s, as recommended by Rosenthal, *Meta-analytic Procedures*, p. 21. However, the mean effect sizes are reported in terms of the untransformed *r*s.
9. Elliot Eisner, "What Justifies Arts Education: What Research Does Not Say," In *Enlightened Advocacy: Implications of Research for Arts Education Policy and Practice*, The 1999 Charles Fowler Colloquium on Innovation in Arts Education, ed. M. McCarthy (College Park: University of Maryland, 2000).

10. For an eloquent statement by an independent school on the importance of the arts, see William Oates, "The Annual Report of the Rector: St. Paul's School 1978-1979," St. Paul's School, Concord, New Hampshire.

Acknowledgments. This research was supported by the Bauman Foundation. We thank John Landrum Bryant of the Bauman Foundation for his support and guidance. We thank Ron Butzlaff for statistical advice and for providing us with the program to perform meta-analyses. We also thank Hiram Brownell for statistical advice. Requests for reprints should be sent to Ellen Winner, Department of Psychology, Boston College, Chestnut Hill, MA 02147 or email ewinner@mediaone.net.