Can Music Be Used to Teach Reading?

RON BUTZLAFF

There are several possible reasons to hypothesize that instruction in music may help children acquire reading skills. First, music and written text both involve formal written notation which must be read from left to right. In both cases, the written code maps onto a specific sound. Perhaps practice in reading music notation makes the reading of linguistic notation an easier task.[1]

Second, skill in reading requires a sensitivity to phonological distinctions, and skill in music listening requires a sensitivity to tonal distinctions. Perhaps experience in listening to music trains a general kind of auditory sensitivity that is as useful in listening to music as it is in perceiving phonological distinctions.[2]

Third, when students learn the lyrics of songs they may engage in reading written text. The lyrics of songs are often repetitive and hence predictable. It is possible that experience in reading such predictable text may train reading skills.[3]

Finally, a motivational argument can also be made. When students form part of a music group, such as a school orchestra or band, they must learn to work together. They also learn that if they do not do their part, the entire group suffers. Hence, there is pressure on instrumental students to be responsible and to work hard. Perhaps experience working in an instrumental group instills a sense of personal responsibility which in turn leads to heightened academic responsibility and performance.[4]

These are the major kinds of hypotheses that have been advanced to justify the study of the relationship between music education and literacy skills. In what follows, I review the empirical literature testing the claim that there is indeed an association between instruction in music (usually school-based) and performance in reading (as measured by reading test scores or by general tests of verbal aptitude). Unfortunately, the available studies do not allow me to tease apart the various possible hypotheses about why music might (if it does) enhance reading.

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

Literature Search

The REAP researchers searched seven electronic data bases from their inception through 1998: Arts and Humanities Index (1988-1998), Dissertation Abstracts International (1950-1998), Educational Resource Information Clearinghouse (1950-1998), Language Linguistics Behavioral Abstracts (1973-1998), MedLine (1966-1998), PsychLit/PsychINFO (1984-1998), and Social Science Index (1988-1998). The search term music was combined with the following search strings: (instruct or train) and (educate or learn or cognition or achieve or intelligence or IQ) and (measure or outcome or effect or evaluation) and (read). In addition, they conducted hand-searches of 41 journals from 1950 to 1998 (listed in Table 1 of the introductory paper in this issue) that publish articles in education, development, and the arts. They checked the bibliographies of all identified articles and sent requests to over 200 arts education researchers for unpublished data or manuscripts not yet published (for which they received a modest rate of return). This search produced a total of ninety-four articles or books that were reviewed for inclusion in this analysis.

I retained for meta-analysis only studies that met the following three criteria: a standardized measure of reading ability was used as the dependent variable; a test of reading followed music "instruction;"[5] statistical information was sufficient to allow for the calculation of an effect size. Studies that randomly assigned children to music vs. control conditions, and that assessed reading ability before and after exposure to music, were classified as experimental; those that did not randomly assign children to conditions and that had no pretest of reading ability were classified as correlational. Six experimental and twenty-five correlational studies were identified and submitted to separate meta-analyses.[6] I first discuss the findings from the correlational studies to determine whether some kind of association between music and reading exists. As will be shown below, the analysis allowed the conclusion that there is indeed such an association. I then turn to the experimental studies to determine whether the relationship between music and reading can be said to be a causal one in which music instruction/exposure leads to enhanced reading ability. As will be shown, no such causal conclusion can be drawn.

Correlational Studies

Table 1 lists the 24 correlational studies included, along with publication date, sample size, effect size, type of music experience the participants received, and the reading outcome measure used.[7] In all of these studies, reading performance by students with some music experience was compared to reading performance by students without music experience. Ten of the studies consist of data provided by the College Board comparing verbal Scholastic Assessment Test scores obtained by students with at least one high school music course to scores obtained by students with no arts courses of any kind in high school. In most of the studies, I make the assumption that the music experience was voluntarily chosen. For example, the College Board studies all assess the verbal performance of students who elect to study music in high school. In all of the studies, no pretest data on reading performance are available; all that is supplied is information on reading performance of students who have already had music instruction. Hence, there is no way to determine whether reading scores improved as a consequence of music study or whether reading scores were high prior to music study.

As the measure of effect size, I have followed the suggestions of Rosenthal[8] and used the Pearson Correlation Coefficient *r*, a statistic which is readily understood by most researchers and readily interpreted. In order to account for the nonnormal distribution of *r*, all of the calculations in the analyses described below were performed on Fisher transformed *r*s which are signified by Z_r in the table.[9] For ease of interpretation, the results reported below are given in untransformed values of *r*.

Effect sizes are shown in a stem and leaf display in Table 2. The mean effect size *r* was $r = .17$, with a weighted mean $r = .19$. The measure of 'robustness' suggested by Robert Rosenthal[10] yielded .94.[11] This value suggests that there is still considerable variability in the effect sizes in this sample, since this statistic is essentially a ratio of the mean effect size divided by the variability around that mean. The combined probability of these studies produced a Stouffer's Z = 301.38, $p <.0001$. This result is highly significant and allows clear rejection of the null hypothesis of no relationship between music and reading. Further support for the existence of a music-reading association is that four-fifths (80%) of these correlational studies had positive effect sizes, with magnitudes ranging from a minimum $r = -.19$ up to a maximum $r = .65$, with a median $r = .17$. In addition, the *t*-test of the mean *Zr* was significant, with $t = 4.2$, $df = 23$, $p <.001$, indicating that the mean effect size was significantly greater than zero. This test is much like a one sample *t*-test in which the null hypothesis examines whether the mean effect size is significantly different from zero. The confidence intervals around the mean *r* at the 95% level ranged from $r = .09$ to $r = .24$. At the 99% level, the confidence interval was $r = .07$ to $r = .27$. Hence, in neither case did the confidence interval span zero. This demonstrates that given another sample of 24 similar studies, we would expect to find that the mean *r* was above zero. The test for the heterogeneity of effect sizes was profoundly significant, resulting in a $\chi^2 = 7671$, $df = 23$, $p < .001$, which means that the effect sizes are not normally distributed.

The file drawer statistic[12] reveals the number of studies together averaging null results (that is, $p = .50$) that would be required to bring the Stouffer's

Table 1: The Twenty-Four Correlational Studies Included in Meta-analysis 1

Author (s)/ Yr.	**Sample Size (*N*)**	**Effect Size *r***	***Z* (*p*)**	**Music type or Music program**	**Reading Test**
College Board (1988)	648,144	.16	125.76 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1989)	587,331	.16	125.98 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1990)	549,849	.17	127.07 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1991)	551,253	.18	136.28 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1992)	545,746	.19	138.42 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1994)	546,812	.21	151.96 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1995)	561,125	.21	159.29 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1996)	568,072	.22	164.75 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1997)	581,642	.22	167.50 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
College Board (1998)	592,308	.22	167.98 (*p*<.0001)	At least one high school course in instrumental or vocal music.	Verbal Scholastic Assessment Test
Engdahl (1994)	598	-.02	-.26 (*p*=.60)	Instrumental Music Program	Comprehensive Test of Basic Skills
Friedman (1959) Fifth Graders	152	-.19	-2.05 (*p*=.02)	Instrumental Music Students	Stanford Achievement Test
Friedman (1959) Sixth Graders	102	.16	1.29 (*p*=.09)	Instrumental Music Students	Stanford Achievement Test
Groff (1963)	460	.02	.35 (*p*=.36)	Instrumental Music Students	Iowa Test of Basic Skills

Author (s)/ Yr.	Sample Size (*N*)	Effect Size *r*	*Z* (*p*)	Music type or Music program	Reading Test
Kvet (1985) District A	17	-.08	-.68 (*p*=.75)	Instrumental Music Students	California Achievement Test
Kvet (1985) District B	42	-.05	-.61 (*p*=.72)	Instrumental music students	Stanford Achievement Test
Kvet (1985) District C	71	.14	.65 (*p*=.26)	Instrumental music students	California Achievement Test
Kvet (1985) District D	45	.18	.68 (*p*=.25)	Instrumental music students	Metropolitan Achievement Test & California Achievement Test
Lamar (1989) 1st Graders - Music Specialists	35	.44	2.41 (*p*=.008)	Students studying music taught by music specialists in school	Stanford Achievement Test
Lamar (1989) 1st Graders - Classroom Teachers	35	.37	1.90 (*p*=.03)	Students studying music taught by classroom teachers in school	Stanford Achievement Test
Lamar (1989) 4th Graders - Music Specialists	35	.65	4.08 (*p*=<.0001)	Students studying music taught by music specialists in school	Stanford Achievement Test
Lamar (1989) 4th Graders - Classroom Teachers	35	.26	1.12 (*p*=.13)	Students studying music taught by classroom teachers in school	Stanford Achievement Test
McCarthy (1992)	957	.10	3.09 (*p*=.001)	Orchestra students	SRA Reading
Weeden (1971)	47	-.06	-.49 (*p*=.69)	Suzuki violin program students	Stanford Achievement Test

Table 2: Stem and Leaf Display of 24 Effect Size *r*s from Correlational Studies

Stem	Leaf
+.6	5
+.5	
+.4	4
+.3	7
+.2	1, 1, 2, 2, 2, 6
+.1	0, 4, 6, 6, 6, 7, 8, 8, 9
+.0	2
-.0	2, 5, 6, 8
-.1	9

Z-test to the just barely significant level of $p = .05$. The file drawer calculation indicates that 805,587 studies averaging null results would need to be found to render the *Z*-test barely significant ($p = .05$).

This analysis demonstrates that there is indeed a strong and reliable association between the study of music and performance on standardized reading/verbal tests. However, correlational studies cannot explain what underlies this association. For example, it is possible that students who are already strong in reading choose to study music; it is possible that students who are interested in music are also interested in reading because they come from families which value both music and reading; or it is possible that a causal relationship exists, such that either music instruction transfers to reading achievement or the reverse. For a test of the directional and causal hypothesis that instruction in music leads to heightened achievement in reading, an examination of experimental studies is required.

Experimental Studies

Table 3 lists the experimental studies included, along with publication date, sample size, effect size, type of treatment condition, and dependent variable measure. Effect sizes are shown in a stem and leaf display in Table 4.[13] The mean effect size *r* for the experimental studies was $r = .18$, with a weighted effect size of $r = .11$ (weighted by sample size). The Stouffer's *Z*-test for combined probabilities proved significant, $Z = 2.38$, $p = .009$. The fact that this test was significant might lead us to believe that we should reject the null hypothesis that there is no relationship between music and reading. However, there are several reasons why we cannot reject this null hypothesis, as explained below.

First, let us examine the confidence intervals around the mean *r*. At the 95% level, the confidence interval is $r = -.21$ to $+.52$ as the upper limit; at the 99% level, $r = -.41$ to $+.67$ as the upper limit. Both of these intervals span zero. In addition, only three of the six studies have positive effect sizes. The 'robustness' value is .48, suggesting that there is considerable variability around this small effect size, and this in turn means that the mean effect size, while significant statistically, is not robust. The wide range of effect

Table 3: The Six Experimental Studies Included in Meta-Analysis 2

Author (s) Yr.	**Sample Size (*N*)**	**Effect Size (*r*)**	*Z* (*p*)	**Type of Treatment**	**Reading Test**
Douglas & Willats (1994)	12	.64	2.0 (*p*=.02)	Music Program (instrumental)	Schonell Reading Test
Fetzer (1994)	30	.57	3.07 (*p*=.001)	Singing Songs	Test of Early Reading Ability-2
Kelly (1981)	42	.06	-.51 (*p*=.70)	Music Instruction	Botel Milestones Reading Test
Olanoff & Kirschner (1969)	46	.00 [a]	.00 (*p*=.50)	Music Training - Special Program	Metropolitan Achievement Test - Reading Comprehension
Roberts (1978)	33	.00 [b]	.00 (*p*=.50)	Note reading, keyboard instrument	Basic Reading Rate Scale
Roskam (1979)	24	-.34	1.28 (*p*=.10)	Music Therapy	Peabody Individual Achievement Test - Reading Comprehension

Notes:
[a] reported as "no significant difference," entered as $r = .00$, $p = .50$
[b] reported as "no significant gain," entered as $r = .00$, $p = .50$

Table 4: Stem and Leaf Display of 6 Effect Size *r*s from Experimental Studies

Stem	Leaf
+.7	
+.6	4
+.5	7
+.4	
+.3	
+.2	
+.1	
+.0	0, 0. 6
-.0	
-.1	
-.2	
-.3	4

sizes also demonstrates this large variability, with the smallest $r = -.34$ and the largest $r = +.64$, a median effect size of $r = .03$, and a standard deviation of the effect size *r*'s of $SD = .38$. The chi-square test of the heterogeneity of the effect sizes was also significant, $\chi^2 = 17.94$, $df = 5$, $p = .003$, indicating that the effect sizes from the studies in this sample are not normally distributed, and that there is significant heterogeneity in this sample of effect sizes.

Further evidence for the nonrobustness of the mean effect size comes from the result of the *t*-test of the mean *Zr*. For this sample, the *t*-test of the mean *Zr* is $t = 1.06$, $df = 5$, $p = .34$, which is evidence that the mean effect size *r* is not significantly different from zero.

A file drawer analysis further supports this interpretation. I found that only seven studies squirreled away in researchers' file drawers and averaging null results would be needed to render the results barely significant. Given that unpublished studies in researchers' file drawers are more likely to have nonsignificant results than are published studies, it is not improbable that seven such unpublished, nonsignificant studies actually exist.

In order to try to account for the heterogeneity of effect sizes found, I performed a linear contrast analysis to examine the hypothesis that the magnitude of effect sizes increases over time. The reasoning here is that in more recent studies, experimenters often explicitly set out to show that music had a positive impact on students' academic performance, whereas in earlier years the researchers were merely trying to demonstrate that allowing students to attend "pull-out" music programs in place of regular class time would not decrease academic performance. Two different experimenter expectancies are suggested by these differing hypotheses, and I wondered if the effect sizes varied in the same direction as these two expectancies. The contrast *r* for publication year was $r = .81$, and this was significant, $Z = 3.45$, $p = .0003$. The magnitude and direction of effect sizes significantly changed with publication year, from negative to positive.

Discussion

These two meta-analyses present two very different pictures. The meta-analysis of the correlational studies shows that students studying music do in fact have significantly higher scores on standardized reading tests (or on the verbal portion of the Scholastic Assessment Test). The mean effect size found, though small, was more robust than that found for the experimental studies, and a considerable number of file drawer studies would be needed to alter this finding. Of course, however, correlational studies allow no conclusion about causality. While the correlational results are consistent with the interpretation that music study enhances reading ability, the results are equally consistent with other noncausal interpretations. For instance, students who score well on reading tests may for some reason choose to pursue music; they may be better equipped for some reason to learn music; or they may read more and their reading experience may enhance their musical interests. Neither the existence nor the direction of causality (if there is causality) can be established in these studies.

The experimental studies, which are designed to test the hypothesis that music study enhances (or causes) reading improvement, yielded no reliable effect. A very small number of file drawer studies could overturn the significance of the result. In addition, there is considerable variation in the effect sizes, indicating that the overall finding is not stable. The correlational studies show this same heterogeneity of effect sizes. In both populations, there are probably other effects not accounted for by the researchers that account for this heterogeneity.

The effect sizes varied widely in the six experimental articles. A brief discussion of the individual articles may help explain some of this variation. The study by Douglas and Willats yielded a significant and large effect size $r = .64$, but this study also found a significant and large interaction ($F(1,10) = 7$, $p = .02$, $r = .64$). This interaction occurred because the control group scores were lowered at the same magnitude as the music group scores were elevated. The authors do not describe who taught the music group, but it is likely that the study was not a double-blind experiment. The researchers probably knew who was in each group and had contact with both the treatment and control groups. Thus, experimenter expectancy effects cannot be ruled out. In addition, the researchers chose subjects with an eye for those students who "might benefit from extra help."[14] Other research has shown that low-achieving students are often the ones who most benefit from teacher expectancy effects.[15]

A similar explanation could account for the large effect found in Fetzer's study. Here the music group was taught music by the researcher, while the control group was instructed by another music teacher. Thus, experimenter expectancy effects could have helped the music group.[16] In addition, there was a fairly large subject loss from the control group (eight out of twenty, or 40%) while in the treatment group, only one subject out of nineteen, or 5%, dropped out. This suggests that something different was occurring in the two groups. Finally, the music group was given more attention than the control group, having been videotaped as a group on three occasions. This extra attention alone is a confounded variable and could possibly account for the positive effect size found in this study.

The study by Roskam had a negative effect size, based upon the scores for music and reading comprehension. However, two other scores were also reported, one for spelling and another for word recognition. These effect sizes were positive ($r = .34$ and $r = .37$, respectively), and one could argue that the average of the three effect sizes ($r = .12$) should have been used in the meta-analytic calculations. However, this average would actually have rendered the results of the meta-analysis less stable and less reliable: only one study would have been needed to bring the Stouffer's Z to barely significant, and the confidence intervals would still clearly include zero (-.05 to +.51 at the 95% level). Because spelling and word recognition are not equivalent to actual reading comprehension, I felt it was best to include only the effect size from reading comprehension. It should also be noted that the author reported that the music group demonstrated much more "severe behavioral difficulties"[17] than did the control group. Thus, although the author states that a larger sample might have yielded a more positive result, this is not likely.

These meta-analyses of studies assessing the relationship between music or music education and reading test scores show that the somewhat significant relationship between these two variables in the experimental studies is neither large, robust, nor reliable. However, only six relevant experimental studies were found, a very small number. In addition, two were assigned an effect size of zero because the author reported no significant difference between groups but did not report statistics that made possible the computation of an exact effect size. This is a conservative solution to this problem, and it is possible that these two studies had positive, though small effect sizes. Finally, the wide variability in effect sizes suggests that further research is needed: as shown in Table 3, two studies were associated with large positive effects; three with minimal or no effects; and one with a negative effect. The fact that two experimental studies did produce large effect sizes suggests that further exploration of this question is merited.

The contrast on publication year was performed to examine possible expectancy effects in the data. Researchers carrying out the more recent studies are likely to be expecting a positive relationship between music and reading, since this kind of relationship has been touted more and more by arts advocates as a justification for music programs in schools. I found that the effect sizes increased from generally negative or negligible sizes to larger, positive magnitudes as the publication years go by. This suggestion of expectancy effects calls for a more stringent research methodology by the experimental researchers before it can be adequately addressed.

It is worth noting that there exists a large body of studies that address a different but related question to the ones addressed here. Many researchers have examined whether music *interferes with* or *enhances* academic performance. In these studies, students hear music at the same time as they read. The effects of music listening on reading have not been adequately summarized. A meta-analysis of these studies would help to clarify another possible relationship between music and reading.

NOTES

1. For these suggestions, see L. L. Kelly, "A Combined Experimental and Descriptive Study of the Effect of Music on Reading and Language" (Ph.D. diss., University of Pennsylvania, 1981), and D. L. Roberts, "An Experimental Study of the Relationship between Musical Note-Reading And Language Reading" (Ph.D. diss., University of Missouri, 1978).
2. Researchers have argued that a structured program in music may help children develop a "multi-sensory awareness and response to sounds," p. 99, from S. Douglas and P. Willats, "The Relationship between Musical Ability and Literacy

Skills," *Journal of Research in Reading* 17, no. 2 (1994): 99. It has also been suggested that music may provide cognitive strategies for "dealing with sound material," L. L. Kelly, "A Combined Experimental and Descriptive Study of the Effect of Music on Reading and Language." (Ph.D. diss., University of Pennsylvania, 1981): p. 22. See also K. Roskam, "Music Therapy as an Aid for Increasing Auditory Awareness and Improving Reading Skill," *Journal of Music Therapy* 16, no. 1 (1979): 31-42. Roskam used music therapy in order to try to increase auditory awareness in individuals with reading problems.

3. For this argument, see L. Fetzer, "Facilitating Print Awareness and Literacy Development with Familiar Children's Songs" (Ph.D. diss., East Texas University, 1994).
4. This motivational argument is put forth by Olanoff and Kirschner and by McCarthy.
5. Thus, studies in which students read while listening to music were excluded. These were considered distraction/interference studies testing the hypothesis that music interferes with reading, rather than training studies testing the hypothesis that music skills help to improve reading ability. In addition, studies in which students were trained to read musical notation, but received no other music instruction, were excluded.
6. Full references of these studies are listed in Works Cited. Studies included in the meta-analyses are asterisked.
7. Table 1 includes three studies (by Friedman, Kvet, and Lamar) from which I could properly calculate more than one effect size because the authors independently sampled from more than one grade or type of student (Friedman, Lamar) or from more than one school district (Kvet). From Freedman's work I calculated two independent effect sizes, and from Kvet and Lamar's studies I calculated four independent effect sizes each.
8. Robert Rosenthal,. *Meta-analytic Procedures for Social Science Research*, (Newbury Park, Calif.: Sage Publications, 1984).
9. Ibid.
10. Ibid.
11. To compute robustness, one takes the ratio of the mean effect size divided by a measure of the variability around that mean, which is the standard deviation of the effect sizes.
12. Robert Rosenthal and Ralph Rosnow, *Essentials of Behavioral Research: Methods and Data Analysis* (New York: McGraw Hill, 1991).
13. One relevant study that was identified could not be used: Irving Hurwitz, Peter Wolff, Barrie Bortnick, and Klara Kokas, "Nonmusical Effects of the Kodály Music Curriculum in Primary Grade Children," *Journal of Learning Disabilities* 8 no. 3 (1975): 167-74. This study reported a significant increase in reading scores on the Metropolitan Achievement Test for children receiving Kodály musical instruction versus those children from the same school who had not. Though the effect as reported would be in the positive direction, the authors do not report enough statistical information to allow for the calculation of a proper effect size for this study.
14. Douglass and Willatts, p. 106.
15. Madon, Jussim, and Eccles, 1997.
16. Rosenthal and Jacobson, 1968.
17. Roskam, "Music Therapy as an Aid,"p. 41.

Works Cited

Articles preceded by an asterisk were included in one of the two meta-analyses.

Cohen , J., *Statistical Power Analysis for the Behavioral Sciences* (New York: Academic Press, 1977).

College Bound Seniors Profile of SAT and Achievement Test Takers. College Board, 1987-1997.

*Douglas, S., and P. Willats, "The Relationship between Musical Ability and Literary Skills." *Journal of Research in Reading* 17, no. 2 (1994): 99-107.

*Engdahl, Pherbia M., "The Effect of Pull-out Programs on the Academic Achievement of Sixth-Grade Students in South Bend, Indiana" (Doctoral diss., Andrew University, 1994).

Fetzer, L., "Facilitating Print Awareness and Literacy Development with Familiar Children's Songs" (Doctoral diss., East Texas University, 1994).

Friedman, Bernard, "An Evaluation of the Achievement in Reading and Arithmetic of Pupils in Elementary School Instrumental Music Classes" (Doctoral diss., New York University, 1959).

*Groff, F. H., "Effect on Academic Achievement of Excusing Elementary School Pupils from Classes to Study Instrumental Music" (Doctoral diss., University of Connecticut, 1963).

Hurwitz, Irving, Peter Wolff, Barrie Bortnick, and Klara Kokas, "Nonmusical Effects of the Kodály Music Curriculum in Primary Grade Children." *Journal of Learning Disabilities* 8, no. 3 (1975):167-74.

*Kelly, L. L., "A Combined Experimental and Descriptive Study of the Effect of Music on Reading and Language" (Doctoral diss., University of Pennsylvania, 1981).

*Kvet, Edward J., "Excusing Elementary School Students from Regular Classroom Activities for the Study of Instrumental Music: The Effect on Sixth-Grade Reading, Language, and Mathematics Achievement." *Journal of Research in Music Education* 32 (1985): 45-54.

*Lamar, H. B., "An Examination of the Congruency of Music Aptitude Scores and Mathematics and Reading Achievement Scores of Elementary Children" (Doctoral diss., University of Southern Mississipi, 1989).

Madon, S., L. Jussim, and J. Eccles, "In Search of the Powerful Self-fulfilling Prophecy." *Journal of Personality and Social Psychology* 72, no. 4 (1997): 791-809.

*McCarthy, Kevin J., "Music Performance Group Membership and Academic Success: A Descriptive Study of One 4-year High School" (Paper presented at the Colorado Music Educators Association, 1992).

*Olanoff, M. and Kirschner, L. C., "Musical Ability Utilization Project," U.S. Department of Health, Education and Welfare, Final Report, Project No. 2600 (1969).

*Roberts, D. L., "An Experimental Study of the Relationship between Musical Note-Reading and Language Reading" (Doctoral diss., University of Missouri, 1978).

Rosenthal, Robert, *Meta-Analytic Procedures for Social Science Research* (Newbury Park, Calif.: Sage Publications, 1984).

Rosenthal, R., and L. Jacobson, *Pymalion in the Classroom: Teacher Expectations and Pupils' Intellectual Development* (New York: Holt, Rinehart and Winston, 1968).

Rosenthal, Robert, and Ralph L. Rosnow, *Essentials of Behavioral Research: Methods and Data Analysis* (New York: McGraw-Hill, 1991).

Roskam, K., "Music Therapy as an Aid for Increasing Auditory Awareness and Improving Reading Skill." *Journal of Music Therapy* 16, no. 1 (1979): 31-42.

*Weeden, Robert E., "A Comparison of the Academic Achievement in Reading and Mathematics of Negro Children Whose Parents Are Interested, Not Interested, or Involved in a Program of Suzuki Violin" (Doctoral diss., North Texas State University, 1971).

Acknowledgments. This research was supported by the Bauman Foundation, I thank John Landrum Bryant of the Bauman Foundation for his support and guidance. I thank Kristin Burger, Lois Hetland, Brian Moss, Leah Okimoto, and Daniel Schneider for identifying all potentially relevant studies and acquiring the studies included in this report. I thank Joanna Holtzman for initial summary and coding of studies. Requests for reprints should be sent to *butzlaff@oakland.edu*.