The Other Mozart Effect: An Open Letter to Music Educators

Robert A. Duke

Update: Applications of Research in Music Education 2000; 19; 9
DOI: 10.1177/875512330001900103

The online version of this article can be found at:
http://upd.sagepub.com

Published by:
SAGE
http://www.sagepub.com

On behalf of:
MENC: The National Association for Music Education

Additional services and information for Update: Applications of Research in Music Education can be found at:

Email Alerts: http://upd.sagepub.com/cgi/alerts

Subscriptions: http://upd.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://upd.sagepub.com/cgi/content/refs/19/1/9
THE OTHER MOZART EFFECT: AN OPEN LETTER TO MUSIC EDUCATORS

ROBERT A. DUKE
University of Texas at Austin

There is a great deal of attention being paid to music in schools lately, and not all of it is positive attention. As subjects compete for time and money in school curricula, questions are asked by curriculum planners, school boards, principals, and parents about the relative value of the arts and whether it is justified to invest the time and money required to sustain arts education in schools when students' performance on math, science, and language arts components of standardized measures (e.g., TAAS, TIMSS, NAEP) is not what all of us would like it to be. Many of these questions about the relative value of the arts are directed at us. And if this isn't bad enough, at the same time that we're supposed to be pondering and responding to questions about the value of what we do, we've got a show to chart for next week, and a holiday program to put together, and I can't find any of the sharp keys for my barred instruments.

I feel compelled to write about this issue now because, in the midst of our dealing with all that educators have to deal with, I think it's not only possible, but even likely, that we may lose sight of what music education is or should be about, and in some ways, at least as far as the popular press is concerned, that may have already happened. In the morass of the great debates in education—of standards and assessments, and whole language and phonics, and block schedules, school choice, and competition—it becomes relatively easy to lose sight of the underlying purpose of what we're doing. Those seemingly ridiculous questions that we pose to freshman education majors—What does it mean to teach? What is the value of knowledge? Why are you here? (questions that no freshman has the vaguest notion of how to answer)—actually do mean something. And even though it may seem like inexpedient, superfluous, ivory-tower mumbo jumbo, time spent thinking about these questions is still very important.

Much of the recent attention to music in our professional journals and in the popular press begins with phrases like, "Research has shown" and "Researchers have found," followed by enticing results about changes in peoples' brains caused by exposure to music and participation in music activities. Having conducted research in music for over 15 years, I'm both gratified and amazed that so much attention is being lavished on an aspect of scholarly inquiry that had received so little attention in the past. But even as opening clauses like "Research has shown" garner increasing credibility for the claims that follow, it seems especially important that all of us continue to remind ourselves of the stringent methodology of systematic inquiry and of what is necessary before we can infer that what's been "shown" is actually true, or meaningful, or important.

Over the past decade or so, many articles have been written explaining the putative benefits of music participation, music study, and music listening. Some of these articles report systematic observations that are consonant with what many of us teachers witness nearly every day in our professional lives. Other articles report subtle effects that are less obvious to the naive observer and that require sensitive measurements made in carefully controlled settings.

An Informal Ethnography

I report below a series of my own observations, recorded over the past year, of children and adults involved in various music activities. I would imagine that the events I describe below are quite similar to those experienced by many teachers and other, generally observant people.

Observation 1. August, 1998. Texas. Groups of students and teachers march around on practice fields...
and parking lots in searing heat, playing wind and percussion instruments for hours and hours each week, though school has not yet opened. Throughout this process, the children develop a sense of pride in what they do, a sense of belonging to a group, a sense that they are contributing to a common goal that extends beyond themselves, and a sense of accomplishment in creating a musical product that elicits an emotional effect in the hundreds of football fans who witness their performances on the field and in the stands. Throughout this process, they are engaged in an important, authentic, substantive, meaningful artistic experience that connects them with their culture. What explains these students' remarkable investment of time, effort, energy, and disciplined commitment to refining physical and musical skills? I have labeled this phenomenon the "Sousa Effect."

Observation 2. February, 1999. Florida. A group of children work day and night to put on a production of South Pacific. In between rehearsing their lines with the drama coach and rehearsing their songs with the music teacher, they're painting sets, sewing costumes, and making posters to advertise their upcoming performances. They demonstrate musical and theatrical skills that have been refined to a high degree; they, too, are engaged in an important, authentic, substantive, meaningful artistic experience that connects them with their culture. What could possibly explain these students' singular focus of attention, consistency of effort and cooperation, and their ability to bring together a complex constellation of skills in a culminating artistic product? I call this the "Rodgers and Hammerstein Effect."

Observation 3. April, 1999. Wisconsin. There are rooms full of fourth graders playing soprano recorders. As they work diligently to produce a consistent tone and to play with precise rhythm and in tune, all of the students listen carefully, concentrate, and think hard as they translate the notation on the board into fingerings and breath and, eventually, sounds that emerge from their instruments as musical expression. They're engaged in an important, authentic, substantive, meaningful artistic experience that connects them with their culture. What explains this attentiveness and diligence among children who are otherwise seemingly unable to demonstrate such sustained effort and obtain such precise and beautiful results? I call this the "Hot Crossed Buns Effect."

Observation 4. June, 1999. California. In a park, a group of teenagers sit listening to music from a jam-box, moving their feet and singing along with a song that speaks to them in a way that many adults don't quite understand. They are compelled by the music; they are focused on the lyrics; and they are moved by the message the song conveys. They are engaged in an important, authentic, substantive, meaningful artistic experience that connects them with their culture. What could explain these children's investment of money (their own and their parents') in purchasing CD's and concert tickets and their investment of time and focused attention in such an activity? I call this the "Boyz II Men Effect."

Observation 5. September, 1999. New York. A young child sits at the piano, as he does on most weekdays, practicing exercises and pieces for his lesson. He struggles mightily to perfect a passage from a Mozart sonata with an awkward fingering, playing the figure slowly at first, then gradually faster until he performs it correctly in several consecutive repetitions. He's more attentive, and his efforts are more concentrated than in just about any other intellectual or physical activity in which he participates. When he finally is able to negotiate the passage, he plays through the entire movement with a musical character that reflects his feeling of pleasure, having recognized that his efforts have led to a tangible accomplishment. All the while, he is engaged in an important, authentic, substantive, meaningful artistic experience that connects him with his culture. What explains this child's focus, concentration, extended effort, expression of ineffable emotion, and sense of personal satisfaction and accomplishment? I have labeled this phenomenon the "Other Mozart Effect."

I chose the Other Mozart Effect as the label for Observation 5 above, because, as everyone in the English-speaking world is now well aware, a Mozart Effect has already been discovered, or so it has been reported. In fact, even the term, Mozart Effect, is now a registered trademark. This first Mozart Effect is the very slight increase in a very specific and narrow aspect of cognitive functioning that sometimes occurs after listening to Mozart's Sonata for Two Pianos in D-major, K. 448, a work whose power to affect the human condition has long been recognized, but, until recently, for entirely different reasons than those reported in the popular press. It has even been purported that the Mozart Effect exists across species. Not only do college students' IQ's increase with exposure to the Mozart Sonata as their peers' IQ's are dulled by the sounds of rock-and-roll, but even laboratory rats run mazes faster after having listening to the Second Sonata, while other rats run slower, having listened to the numbing minimalism of Phillip Glass.

How Science Works

If our discipline is to develop an informed perspective about the weight of evidence for the Sousa Effect,
the Rogers and Hammerstein Effect, the Hot Crossed Buns Effect, the Boyz II Men Effect, and both Mozart Effects, then we must consider carefully three issues related to experimental research. The first issue concerns the way that science goes about its business and what the process of systematic inquiry means for establishing the trustworthiness of any hypothesis or theory. After observing phenomena in the natural world, scientists posit hypothetical explanations for why things are the way they are and then make predictions about what will happen in controlled situations when certain variables are present or absent. Good experiments are then designed in a way that affords every conceivable opportunity for the proposed hypothesis to fail, so that the effects observed can be rightfully attributed to the suspected causes.

A hallmark of scientific inquiry is replication, the process of repeating experiments in an effort to demonstrate that identical conditions produce consistent results. The results of one experiment, no matter how dramatic, are insufficient evidence for confirming any hypothesis. Repeated experiments are necessary before one can say with confidence that a hypothesis is “true,” and even after repeatedly obtaining identical results, all hypotheses remain open to continued scrutiny, refinement, and even rejection.

The second important issue in interpreting scientific evidence concerns the meaning of the term statistical significance. The word “significant” in statistical jargon has a very different meaning than in common parlance. In our day-to-day speech, most of us use the word significant to describe things that are important or notable. Statistical significance has an entirely different and much narrower meaning. Statistical significance refers to the likelihood (probability) that an experimental result is attributable to chance or random error. All experiments represent singular instances of observation (samples of all possible instances). In all experimental data (especially data involving groups of people), there is an element of experimental error (which appears as variations among individual data points [e.g., subjects’ test scores]), that cannot be attributed to any identifiable cause. Statistical tests calculate the probability that the results obtained in an experiment could have appeared only as a result of these random variations in the absence of any true experimental effect. If, on the basis of one or more statistical tests, a researcher concludes that his or her results are significant, this conclusion means only that the results are unlikely to be attributable to random variation. Put most simply, “significant” results are those that are probably not mistaken. That’s it. Note that statistical significance says nothing about either the magnitude or the importance of an observed effect. Even statistically significant effects can be so small as to be meaningless.

The third issue of importance concerns the magnitude of experimental effects. The term effect size, as its name suggests, is an empirical measure of the magnitude of an experimental effect. More specifically, effect size describes the extent to which the subjects who are exposed to an experimental treatment (e.g., listening to Mozart) differ from a control group who are not exposed to the same treatment (e.g., sitting in silence). The importance of effect size in evaluating the results of formal experiments and informal observations is immediately apparent. Effect size is the research-jargon answer to the question: “How much difference does it make?” Once research determines that an experimental effect is statistically significant (i.e., not likely a result of random error), the next logical question concerns the size of the effect, and it is the question of effect size that is most closely related to importance.

For example, the two routes that I travel between my home and my office are significantly different from one another in terms of travel time—significantly different in the statistical sense, that is. The time it takes me to get from one place to the other on multiple trips over each route is highly reliable (consistent). Route 1 averages 20 minutes, 18 seconds. Route 2 averages 19 minutes, 53 seconds. Because of the consistency of the timings of these two routes (each one takes about the same amount of time on any given day), the second route is significantly shorter (in the statistical sense) than the first route. But the magnitude of the difference between the routes is only 25 seconds, a difference that is certainly unimportant. Although the difference between the routes is statistically significant, no one but a statistician would call Route 2 significantly shorter than Route 1. Even though the difference is reliable (consistent) and significant (not a result of random variations in travel times day to day), the effect size is very small (the difference is unimportant, its statistical significance notwithstanding).

I know that many nonscientists are frustrated by widely publicized pronouncements about recent discoveries concerning diet and health, for example, only to learn later that the effects of effortful dietary modifications are, at best, minuscule and, at worst, invisible. The recent data on dietary salt is a stunning example. Grand conclusions were effectively disseminated in the popular press about relationships between the intake of dietary sodium and hypertension, high systolic and diastolic blood pressure, and heart attack and stroke. Even though these conclusions were based on carefully conducted research, the leap from a few,
weak statistical relationships to emphatic rules of salt consumption was large and unjustified.

There were, in fact, statistical relationships between sodium intake and the maladies with which salt was said to be associated, but the magnitude of these relationships—the effect size—was vanishingly small in all but a few narrowly defined populations of people. But the pronouncements in the popular press failed to include any of these qualifications. There was a purported relationship between sodium intake and health. Salt is bad. Get rid of salt in your diet. But the available data did not and do not support that conclusion. So how did such a mistake come about? How could results be reported with such confidence when there were not sufficient data to support the conclusions? And, in a broader sense, what is the effect of making such strong pronouncements, only having to retract them later? Well, you know the answer to the last question. Many people start to believe that those health researchers have no idea what they're talking about. They said that this was true, that they were certain, and now they're not so certain. So why pay attention?

Most of us find it convenient to think of many questions in terms of dichotomous answers. Is this true or is that true? Right or wrong? In tune or out of tune? Of course, most questions are not so clearly dichotomized. There is more than just right and wrong; there are degrees of correctness, qualifying conditions, but these take time to explain and understand. Given our general need to decide, we often find either-or decision making rather appealing.

So it seems important, when evaluating experimental results that potentially affect our discipline, to concern ourselves not only with whether an effect is statistically significant (i.e., not attributable to random variation or experimental error) but, more importantly, to concern ourselves with the magnitude of the effect—how much of a difference does it make?

Here's my point: The effects I first described, somewhat facetiously (the Sousa Effect, the Hot Crossed Buns Effect, etc.), are large enough to be observed every day in nearly every city in America. Similar effects are observable in societies across the globe. The effects sizes of these phenomena are large by any measure, even though most of them are not routinely subjected to statistical analyses. The most recently publicized Mozart Effect, in contrast, has been observed in only a small number of published articles, has not been observed in other attempts to replicate these studies, and, when observed, is very narrowly defined and very small in magnitude. In the letter to the editor in Nature (Rauscher, Shaw, & Ky, 1993), for example, which many believe was the article that brought the so-called Mozart Effect to the attention of the media, the difference between the students who listened to the Mozart Sonata and the other students was three points on tests of abstract/spatial reasoning, a very narrow and specific intellectual task. The authors then "translated" that three-point difference into an IQ score difference of 9 points by making the enormous assumption that the tiny effect observed on the spatial test would generalize across the spectrum of intelligence. Two months later, a psychometrician from Great Britain wrote a letter in response to the first report, citing that the group mean difference between the test scores of the Mozart condition and the other conditions was within the expected error of measurement of the tests used in the experiment. That letter never made the papers. Nor did a study that a colleague of mine in Texas conducted with his psychology class, in which he compared listening to Mozart to staring at a computer screen-saver with flashing images. The students who stared at the computer screen obtained the highest scores in that comparison.

What I find curious about all of this is the seemingly disproportionate attention being paid to an effect that, if real (and that question has yet to be settled), is minuscule compared to the other effects that all of us observe as teachers every day. The obvious explanation for the disproportionate attention is that we're talking intelligence here and math and test scores. Yes, but it's three points. Now I realize that discussions of intelligence and math and test scores have a certain cachet in educational and political circles. And, even those who believe that music is, at best, a frill that is peripheral to the educational mission might change their tune once they realize what's happening to students' brains, especially if these effects can be demonstrated with test scores. Or so it is argued.

But it's important that we remember whom we're talking about. These same individuals who are making decisions about budget allocations and schedules are also familiar with cost-benefit analyses. We certainly may get their attention with our promoting the notion that Mozart makes people smarter, but what will they do when they learn the size of the Mozart Effect, even assuming that there is, in fact, a Mozart Effect? Is it really worth the cost of Sousaphones and bassoons and Orff instruments and octavos and uniforms and music teachers (!) for three points on a test that measures a very narrow aspect of intellectual functioning? What if the money and time currently invested in music instruction in schools were redirected toward more time in math classes and more math teachers? It's hard to imagine that we couldn't move scores on a spatial reasoning test, and ultimately scores...
Inferring Cause from Correlation

Now, some might argue that I'm missing the point. No one is saying that the so-called Mozart Effect is the reason for music instruction. It's simply one more piece of evidence in our favor. Let me respond to that position by suggesting that the proportion of attention given to the Mozart Effect and other statistical comparisons that purportedly illustrate the positive effects of music on students' academic performance indicates otherwise.

Graphs showing SAT score comparisons between students enrolled in arts classes and students not enrolled (graphs that were rife in our published literature even before the Mozart Effect was reported) suggest to most readers that being in those classes has somehow resulted in students' earning higher test scores. Although most of the researchers who publish these graphs include somewhere in the text a caveat stating that the correlations reported do not demonstrate causality—they do not prove that music instruction produced the higher scores—these cautions are not very prominently placed. The presentations of this type of information are akin to diet commercials that feature an attractive, svelte woman who speaks into the camera of her happy new life as she is juxtaposed with an unflattering photograph of her enormous former self, while, as required by the FCC, a message is flashed for mere seconds in 5-point font at the bottom of the screen stating, "Your results may vary." They may indeed.

What do these graphs of SAT scores and arts enrollment show? Well, they show that there are lots of smart kids in music classes. So far, so good. What does that mean, exactly? Well, most likely, it means that smart kids like to be in music classes and that smart kids' parents like for their kids to be in music classes. What a positive thing to say about music and music instruction! What do these data demonstrate about the effects of music study on students' performance in school? Nothing, absolutely nothing. If they did, and if we believed them, we'd have to make some serious changes in what music classes we offer. In Texas, for example, we publish the mean SAT scores for students in each all-state performing ensemble (e.g., the choir, orchestra, band, and jazz band), and the scores are consistently differentiated. In 1999, the mean SAT scores for the all-state orchestras are highest, followed in descending order by the jazz ensemble, the symphonic bands, and the choirs. All of the groups' means are several hundred points higher than the SAT national average.

Of course, we publish these scores because they imply, even though we might not say this directly, that being in music makes one a better student, or at least a better test taker. But if we follow this reasoning to its logical conclusion, then we must ask ourselves why we're wasting valuable time and money on choir, band, and music appreciation, when it's clear that orchestra gets the biggest bang for the buck? Besides, the orchestra plays Mozart. Let's get everyone into orchestra. Of course, no reasonable person would argue that point, because it's clear that these comparisons show only which students end up in which classes, not that being in the classes changed their SAT scores. A comparison of SAT scores between students in AP Calculus and Business Math would show differences even more striking than the music comparison, but I doubt that there are many parents or administrators who would not interpret these findings simply as evidence that brighter kids take more advanced math classes, rather than concluding that taking calculus makes you a brighter kid.

What the Data Actually Say

The sidebar lists publications about research related to music and cognition. The published research related to the Mozart Effect is included in the section labeled Effects of music interventions on cognitive abilities. The results of these studies demonstrate the following:

1. The so-called Mozart Effect has not been reliably observed; a number of investigators attempting to replicate the effect have failed to find evidence that music listening results in superior performance on tests of spatial reasoning.

2. The limited evidence for effects of music listening and music instruction is confined to a very narrow and very specific type of cognitive task (even some tests of spatial reasoning fail to record any evidence of a Mozart Effect).

3. The magnitude of the purported Mozart Effect, even when found to be statistically significant, is very small.

4. The changes in scores on tests of spatial reasoning following music listening may be attributable to heightened attention or arousal, effects that may be produced by stimuli other than music.

5. The so-called Mozart Effect, when observed, is not limited to the music of Mozart; for example, one investigator obtained similar results with music performed by Yanni (the "Yanni Effect").

6. Claims that music listening increases performance in any aspect of mathematics, chess play, or
will inevitably lead to the improvements in students’ performance on intelligence tests (as if that mattered) and their performance in school. Such evidence does not yet exist. Recognizing the basic value of arts education, most advocates for music in schools agree that increased test scores, by themselves, provide a weak rationale for music in schools. Yet, many consider the Mozart Effect just one more weapon in our arsenal for music advocacy. Besides, even if the veracity of claims related to the Mozart Effect is questionable (borrowing from my Aunt Hildie’s response to those who doubted her chicken soup’s potential to cure disease), “it couldn’t hurt.” Well, yes, it could, for reasons I think I’ve already explained.

Second, we need to consider carefully the benefits of music in proportion to their effect sizes, which suggest that we restore focus to the effects of music that are vividly and reliably observable on a day-to-day basis in all of our experiences as teachers, parents, and caregivers. Music is a marvelously engaging and rewarding activity. All of the arts are a basic part of human culture and a fundamental aspect of human communication and expression. To teach our children about the arts is to teach them about the culture and society in which they live, while, at the same time, helping them develop sophisticated skills in auditory and visual discrimination, fine motor skills, and a sense of personal accomplishment through active participation in arts activities. And, if one needs a more practical rationale for music study, there is no better activity through which children can observe a tangible relationship between their own efforts and the results their efforts produce (in terms of increased skill, capacity, and expressive potential) than learning to sing or play an instrument. Measured in terms of personal pleasure, accomplishment, attentiveness, skill development, and personal expression, music study obtains results that are reliable (consistent), significant (not a result of random variations), and large (important) in their effect sizes.

Teaching by Example

We are obliged as educators to provide our students with positive models of excellence in all aspects of academic life, including the evaluation and interpretation of empirical evidence, although experimentation and empirical verification are certainly not the source for all our decision making. But when we base certain of our decisions and personal beliefs on reasons other than empirical evidence and research—as it certainly seems fitting to do—then we need to acknowledge that our mode of decision making does not include a need for empirical evidence. And, in those instances in which our view of the world is formulated in the absence of systematic gathering and evaluating evidence, it would be more than disingenuous for us to then cite evidence we happen to encounter that supports our view and ignore evidence that undermines it.

When we do look to science for answers to questions that we believe science can answer, then we need to follow the rules of the game we’ve agreed to play and not only accept the evidence that makes us happy and polishes our biases. We should require evidence that meets accepted standards and withstands careful and repeated scrutiny. The oft-quoted line from T. H. Huxley is instructive: “My business is to teach my aspirations to conform themselves to facts, not to try and make facts harmonize with my aspirations.” We know many things about music and music study that are incontrovertibly true. Other things that we may like to be true are as yet unproven. It is in our own best interest, and in the best interest of our discipline, to stick to what we know and proceed with due caution and deliberation in expanding our knowledge of what music has to offer.

Selected Resources

**Effects of music interventions on cognitive abilities**


Relationships between music and cognitive abilities


Archives of Psychology, 141, 71–84.


Relationships between music and brain chemistry and physiology


Reviews of research, other articles in journals, and the popular press


Note

1. This article is reprinted with permission from The Southwestern Musician Combined with the Texas Music Educator, the official journal of the Texas Music Educators Association, March 2000, Volume 68, No. 8.