Kenneth M. Steele, Ph.D.

The "Mozart Effect":

An Example of the Scientific Method in Operation

Kenneth M. Steele, Ph.D., Appalachian State University

Students have difficulty understanding and appreciating the value of the scientific method in dealing with issues in psychology. Typically, students see conclusions from a study or two about a complex question. The problem is that the student must accept or reject these conclusions because they come from a particular source, whether teacher or textbook. This is reasoning by the method of authority, which was not the method used to obtain the original results and is not how research psychologists think about such results.

Some outcomes may be so universal that we can treat them as "facts." Other findings are equivocal or enigmatic. The scientific method is a process of empirical evaluation of all findings. Research on the Mozart effect exemplifies this process for two reasons. First, the effect is relatively simple to understand. Students do not have to learn much about equipment or deep issues of research design. Second, a sequence of experiments appeared in rapid enough order that students can appreciate the process.

The Original Finding

The original article appeared in Nature (Rauscher, Shaw, & Ky, 1993). It reported that 36 college students showed an increase on spatial reasoning scores from subtests of the Stanford-Binet Scale of Intelligence after listening to a Mozart piano sonata relative to listening to a relaxation tape or silence. The effect occurred only if the subjects were tested immediately. The size of the effect was the equivalent of 8 to 9 IQ points.

The music selection was from the Sonata for Two Pianos in D Major (K. 448). It is lively and emphasizes the virtuosity of the performers. It is not a central piece in the Mozart canon.

The spatial reasoning measures consisted of a pattern analysis task, a multiple-choice matrices task, and a multiple-choice paper-folding and cutting task. Figure 1 shows an example matrices item. The task is to choose the geometric figure from the lower line which should be inserted in the empty cell to complete the pattern. Figure 2 shows an example paper-folding and cutting item. The top row shows a piece of paper undergoing a fold and a pair of cuts, proceeding from left to right. The task is to pick the illustration in the bottom row that represents the paper when it is unfolded.

Rauscher, Shaw, & Ky (1995) reported a replication of their discovery using only the paper-folding and cutting test.

Why did the effect become famous?

The authors contended that this was the first experiment to demonstrate that listening to music caused an increase in spatial reasoning. The issue of cause is important. Many people believe there is a positive correlation between academic success (like high school grade point average) and musical experience (like participation in a band) although the research literature is ambiguous.

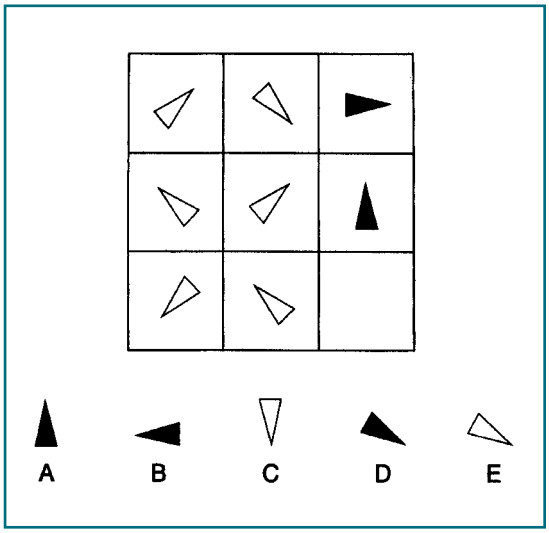

Figure 1. Practice Stanford-Binet Matrices Item. The correct answer is 'B' for the item illustrated here.

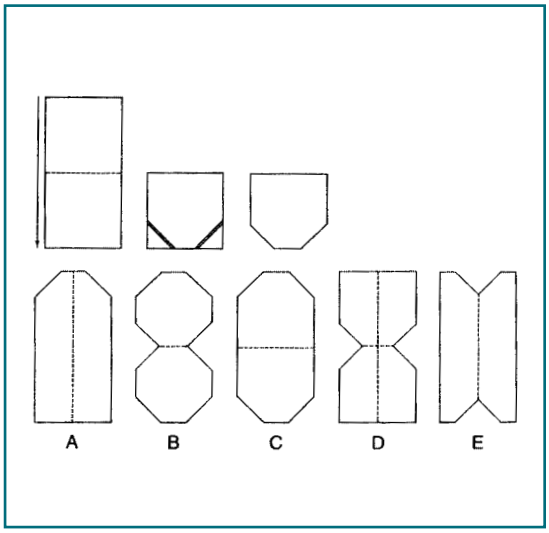

Figure 2. Practice Stanford-Binet Paper Folding and Cutting Item. The correct answer is 'C' for the item illustrated here.

People often assume that a correlation suggests a causal connection. However no firm conclusion can be drawn as to why this relationship exists. For example, students from wealthy homes may have the added time and opportunity to succeed more than average in both mathematics and music.

The authors' causal interpretation was that exposing a person to that specific sonata was the sole factor that explained the increased reasoning scores. There was no explanation in the original article of why the Mozart effect should have occurred. In later publications, Rauscher and Shaw suggested that this particular sonata was activating brain regions required by the spatial reasoning tasks, and that this overlap of activation could be related to a mathematical model of neural activity by Shaw, a physicist by training.

Later, their interpretation was transformed into the global generalization that participation in music activities would produce increases in mathematical performance and that the existence of the Mozart effect demonstrated the academic necessity of music education in the school curriculum.

Mass marketers sold books and CDs to worried parents with the promise that early exposure to the right music would speed intellectual development. Rauscher and Shaw contributed to the frenzy by adding that listening to this sonata could reverse the effects of senile dementia, epileptic seizures, and improve the maze-learning ability of rats.

How to approach research results

A startling claim is made and repeated widely in the press. How is the claim evaluated? First, one must differentiate between the results of the experiment and the interpretation that was applied to the results. The results were that students showed increased scores on specific problems after hearing a portion of piano sonata relative to their scores after listening to a relaxation tape or sitting in silence. The interpretation was that some property of the music, perhaps a certain pattern of notes, increased the activity of brain regions involved in spatial problems and that this increased brain activity produced increased accuracy of solution of visual puzzles.

One can see a large gap between the results and the interpretation. A purpose of the scientific method is to determine whether the gap can be filled successfully by a series of experiments which successively extend the original finding. To do this researchers needed a description of the critical properties of the music, the means by which the music produced its effect, the range of activities that would be affected, and the expected duration of the effect. But the first step is that you need to be able to produce the effect.

Early attempts to produce a Mozart effect

Several experiments appeared after the initial study and reported negative results. Carstens and colleagues (1995) had students listen to the original Mozart sonata and then answer 64 multiple choice items from the Minnesota Paper Form Board Test. Participants viewed two-dimensional parts and selected the figure that indicated the appearance of the final unit when the parts were assembled.

Carstens and colleagues found no difference in performance between the Mozart group and a control group who meditated in silence. No difference, or a null result, is tricky to interpret. The lack of difference could be due to the lack of effect of listening to the sonata or due to other conditions which interfered with the subjects. Carstens and colleagues found that the subjects' SAT scores predicted Form Board scores but the addition of information of whether the subject listened to Mozart changed the size of the prediction score by a trivial amount. The important point is that the lack of a Mozart effect became meaningful in the context of the other expected finding.

A second study worthy of note was by Newman and colleagues (1995). They increased the number of participants so that detecting the effect would be more likely, and they obtained background information on the musical training and preferences of their subjects. Subjects listened to the Mozart sonata, a relaxation tape, or sat in silence and were tested on items from Ravens Progressive Matrices. The task is very similar to that illustrated in Figure 1.

Newman, like Carstens, found no Mozart effect. Additionally, the effect of musical background was not consistent with what would be expected from Mozart-effect advocates. Subjects who had extensive music training (M = 8 years) performed no differently than subjects who had no musical training. Moreover, subjects who indicated a preference for classical music scored significantly worse on the matrices problems compared to those who preferred "other" music.

The Reply of Rauscher and Shaw (1998)

The purpose of Rauscher and Shaw's reply was to explain the difficulties that other researchers were having in producing the Mozart effect. Their major point was that previous experiments did not test the right type of spatial reasoning.

The original 1993 article had reported improvement on the combined measure of the three Stanford-Binet tests because scores on the tasks were well correlated. Rauscher and Shaw (1998) explained that the improvement reported in the 1993 article had occurred only with the paper-folding and cutting task, so the lack of effect observed by Carstens et al.

See Mozart, page 5

Mozart, from page 3

and by Newman et al. with their matrices tasks was not a contradiction.

Second, Rauscher and Shaw suggested that other differences among experiments might interfere with showing the effect. However, they were vague in connecting this suggestion to specific procedures in studies. Finally, they stressed the importance of choice of musical composition. But, five years after the original report, they could be no more specific than "complexly structured music, regardless of style or period."

The Replications by Steele and Colleagues

I had read the Carstens et al. and the Newman et al. experiments in the course of preparing one of my own manuscripts. Their procedures were reasonably similar to Rauscher and Shaw's. Yet there was such a striking difference in outcome. The size of the statistical difference was large in both Rauscher et al. experiments, but was quite miniscule in the Carstens et al. and the Newman et al. experiments. Rauscher and Shaw blamed the problem on the use of the wrong dependent measure.

Perhaps, there was some other aspect of the procedure that explained the difference in results. In this case, the best solution was to replicate the procedure of one of the Rauscher et al. experiments. Like a cook, I would follow her recipe. Having obtained her result then I could vary the recipe to discover what ingredient was causing her bread to rise while everybody else's fell flat.

The Rauscher et al. (1993) report in Nature was very short, and lacked many procedural details. However the 1995 report appeared as a standard length article with the necessary procedural details and had the added bonus of being consistent with the advice of Rauscher and Shaw (1998). The experiment by Steele, Bass, and Crook (1999) replicated the essential details of the 1995 procedure. *(Readers may obtain a copy of this article at <http://www.psych.appstate.edu/Faculty/Steele.htm>)*

Steele, Bass, and Crook (1999) used the same Mozart sonata, the same control condition stimuli, the same paper-folding and cutting task, and the same experimental design. Participants completed a pretest on 16 items. Two days later, subjects were exposed to a treatment condition and then immediately tested on 16 new items. We increased the number of subjects over the Rauscher et al. study to increase the experiment's sensitivity to the effect.

In addition, we used standard variations of design. I was confident that the Mozart effect would be replicated at last, and then the real analysis of discovering the reasons for the contradictory findings would begin. I was shocked with the final results. There was no hint of a Mozart effect. Figure 3 shows the average number of items correct for the three groups when tested immediately after the listening condition.

Using Rauscher's own recipe, our bread was as flat at that reported by Carstens, Newman, and others. A complete list of the studies can be found in a summary article by Christopher Chabris.

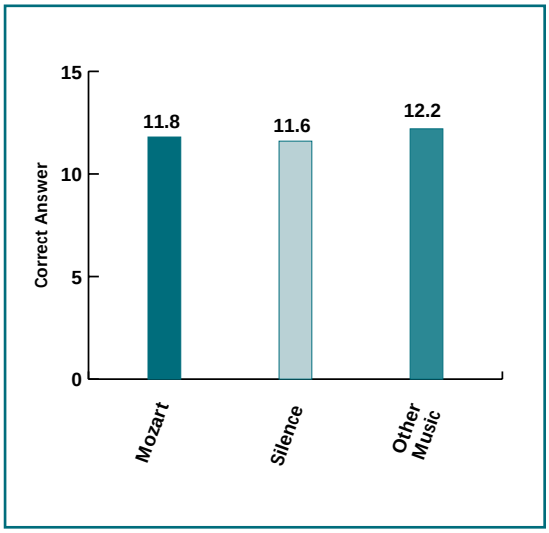

Figure 3. Steele, Bass, & Crook (1999) Results. The figure shows the mean number of Paper Folding and Cutting items answered correctly by subjects after 10 minutes of hearing either the Mozart piano sonata, a period of silence, or listening to other music.

Requiem for the Mozart effect

A requiem is music written to honor the dead. It is odd to speak of a requiem for an effect but it is appropriate in this case. The original report was a startling but isolated claim. The authors provided no testable explanation of why this particular music was endowed with special properties or the nature of those properties. The neurophysiology of the effect and its linkage to mathematical reasoning was fragmented and speculative. The wide-spread endorsements of the effect came from commercial interests and committed advocates instead of the research community.

Yet the effect should be honored because it illustrates the scientific method in operation. The original report was subjected to the same process applied to all scientific claims. The report was followed by a series of studies by independent investigators who sought to verify and understand the effect. Researchers were able to build on the results of earlier investigations and move towards the critical studies, which indicated in this case that the original report could not be verified. The rise and fall of the Mozart effect is a case of effective science in action.

References available upon request from the *PTN*.