

## FAILURE TO CONFIRM THE RAUSCHER AND SHAW DESCRIPTION OF RECOVERY OF THE MOZART EFFECT<sup>1</sup>

KENNETH M. STEELE, JOSHUA D. BROWN, AND JAIMILY A. STOECKER

*Appalachian State University*

*Summary.*—The Mozart effect is an increase in spatial reasoning scores detected immediately after listening to the first movement of a Mozart piano sonata. Rauscher and Shaw (1998) suggested that failure to produce a Mozart effect could arise from carryover effects of a spatial reasoning pretest which may interfere with the effect of listening to Mozart. They cited an unpublished study in which a verbal distractor was inserted between the pretest and listening condition, and the manipulation produced the recovery of a Mozart effect. This experiment attempted to confirm the unpublished study. 206 college students were exposed to one of three sequences, pretest–Verbal distractor material–Mozart, pretest–Mozart–Verbal distractor material, and pretest–Verbal distractor material. An immediate posttest indicated no significant difference on solution of paper folding and cutting items among the three groups. The results do not support Rauscher and Shaw (1998). Our negative results are consistent with prior failures in other laboratories to produce a Mozart effect.

Rauscher, Shaw, and Ky (1993) reported that 36 undergraduates increased their mean spatial reasoning scores the equivalent of 8 to 9 IQ points on portions of the Stanford-Binet Intelligence Scale–Fourth Edition (Thorndike, Hagen, & Sattler, 1986) after listening to the first movement of Mozart's *Sonata for Two Pianos in D Major*, K.448 (hereafter labeled the "Mozart effect"). The Mozart effect was temporary, reportedly disappearing within 10 to 15 minutes. Rauscher, Shaw, and Ky (1995) reported a replication of the Mozart effect, using elaborations of the Stanford-Binet Paper Folding and Cutting task as the dependent measure.

Several researchers have not produced the Mozart effect, despite using a variety of procedures and dependent measures (Carstens, Huskins, & Hounshell, 1995; Kenealy & Monsef, 1994; Newman, Rosenbach, Burns, Latimer, Matocha, & Vogt, 1995; Steele, Ball, & Runk, 1997; Steele, Bass, & Crook, in press; Stough, Kerkin, Bates, & Mangan, 1994), including spatial reasoning tasks from the Stanford-Binet IQ measure (Kenealy & Monsef, 1994; Weeks, 1996). Rauscher and Shaw (1998) have provided an explanation for some of these failures. They caution researchers about two issues.

---

<sup>1</sup>The work reported here was supported by grants from the Cratis D. Williams Graduate School, Appalachian State University. Portions of this experiment were included in a thesis submitted by Joshua D. Brown in partial fulfillment of the requirement for graduation with honors in psychology from ASU and were presented previously at a meeting of the Southeastern Psychological Association. We thank Stan Aeschleman for comments on an earlier manuscript. Address correspondence to Kenneth M. Steele, Department of Psychology, Appalachian State University, Boone, NC 28608 or e-mail (steelekm@appstate.edu).

First, they direct researchers to use the correct dependent measure. This measure must combine both a spatial task and involve a series of transformations over time. According to Rauscher and Shaw, listening to Mozart should not improve performance on measures like the Raven Progressive Matrices (Raven, 1986) or the Backwards Digit Span task because the tasks do not combine spatial imagery and the temporal ordering of spatial components. They describe two tasks as meeting their requirements, the Paper Folding and Cutting task from the Stanford-Binet IQ measure and maze learning.

Also, Rauscher and Shaw (1998) warn researchers about single session within-subject pretest-posttest designs. They suggest that effects of the pretest may continue during the treatment phase, i.e., listening to the Mozart sonata, and this carryover may block the enhancement of spatial reasoning after listening to Mozart and produce no performance improvement during the posttest. In support of this conjecture, they cite an unpublished study by Nguyen, Shaw, and Tran (1996). Nguyen, *et al.* reported finding no significant differences between groups exposed to Mozart or silence until a verbal distractor task was inserted between the pretest and the listening treatment, at which point significant differences between the mean performances of the groups emerged. As the Nguyen, *et al.* (1996) study is unpublished, it is not known whether appropriate control conditions were used.

It was the purpose of the study reported here to investigate whether the insertion of a verbal distractor between a spatial reasoning pretest and presentation of the Mozart D Major sonata would produce the emergence of a Mozart effect. Our dependent measure consisted wholly of the Paper Folding and Cutting items from the Stanford-Binet Intelligence Scale-Fourth Edition (Thorndike, *et al.*, 1986) to meet the dependent measure requirements of Rauscher and Shaw. This study used three treatment conditions. The Verbal-Mozart condition comprised the sequence of a pretest, followed by a verbal distractor, which was then followed by a portion of a Mozart sonata, and, finally, the sonata was followed immediately by the posttest. This condition replicated the sequence of Nguyen, *et al.* (1996) that should result in performance enhancement on the posttest. The Mozart-Verbal condition was comprised by the sequence of the pretest, followed by the Mozart sonata, the verbal distractor, and finally the posttest. This treatment sequence reversed the order of the Mozart sonata and the verbal distractor. There should be no performance enhancement on the posttest for several reasons according to the Rauscher and Shaw (1998) analysis. The effects of solving the Paper Folding and Cutting items in the pretest should carry over and interfere with the hypothesized effect of hearing the Mozart sonata and, further, whatever enhancement effect occurred from listening to Mozart should be further reduced both by the combination of passage of time and listening

to the verbal distractor material. A second control condition was the Verbal condition, which comprised a pretest, verbal distractor, posttest sequence. The purpose of this condition was to assess whether listening to the Mozart sonata produced any enhancement of performance, independent of the presentation sequence.

## METHOD

### *Participants*

Students (133 women and 73 men) from introductory psychology classes participated for course credit.

### *Apparatus*

The Mozart sonata was that used in previous studies by Rauscher, *et al.* (1993, 1995), the first movement of the *Sonata for Two Pianos in D Major* (K.448), performed by M. Perahia and R. Lupu. The verbal distractor was the audio comedy performance *Class Clown*, by George Carlin. The Mozart sonata was 8 min., 24 sec. in duration; the Carlin performance was 16 min., 5 sec. in duration. Audiotapes were created of these selections and were played on a Sony CFD-545 audio unit during the study.

The Paper Folding and Cutting items were taken from the Stanford-Binet Intelligence Scale—Fourth Edition (Thorndike, *et al.*, 1986). This task contains two practice and 18 test items in a graded series. The series of 18 test items were assigned in alternating fashion to create two groups of nine items. Each item was copied onto a 21.5- by 28-cm transparency for use with an overhead projector.

### *Procedure*

Sessions were conducted in the early evening when the building was quiet. Participants were run in groups of 15 to 20 students per session which allowed assurance that the visibility of each projected item was equal for all participants. Participants were informed that they would be participating in a "puzzle experiment." The two sample Paper Folding and Cutting items from the Stanford-Binet IQ measure were used to explain the task. The first sample item was projected by overhead onto a white 3.5-m by 3.5-m screen. The experimenter explained the task using instructions slightly modified and abbreviated from the instructions provided by the Stanford-Binet measure. The second sample item from the Stanford-Binet measure was presented and instructions presented in a similar fashion. Following the opportunity to ask questions, answer sheets were provided and nine test items were projected for up to 1 minute each, with a 5-sec. warning at the end of the 1-min. period. Following exposure to the scheduled stimulus condition (Verbal-Mozart, Mozart-Verbal, or Verbal), participants were immediately tested on a new set of nine items. The two sets of Paper Folding and Cutting items

were used in counterbalanced order across sessions and groups. The distribution of men and women was counterbalanced across groups.

### RESULTS

Table 1 shows mean number of items answered correctly by each group for the pretest and posttest portions of the session. The groups did not significantly differ in performance on the pretest ( $F_{2,203} = 1.08$ , ns). There was an over-all improvement by the groups on the second test ( $F_{1,205} = 24.12$ ,  $p < .001$ ). *Post hoc* tests indicated that the improvement was statistically significant for the Verbal-Mozart group ( $t_{68} = 3.26$ ,  $p = .002$ ) and for the Verbal group ( $t_{68} = 3.81$ ,  $p < .001$ ) but not for the Mozart-Verbal group ( $t_{68} = 1.62$ , ns).

TABLE 1  
NUMBER OF STANFORD-BINET PAPER FOLDING AND CUTTING ITEMS  
ANSWERED CORRECTLY AT PRE- AND POSTTEST BY CONDITION

Treatment Condition	Pretest		Posttest		n
	M	SE	M	SE	
Verbal-Mozart	6.85	0.28	7.48	0.25	67
Mozart-Verbal	6.56	0.27	6.91	0.24	70
Verbal	6.23	0.28	7.04	0.24	69

There was no significant effect for the treatment on posttest performance ( $F_{2,203} = 1.45$ , ns) or on the interaction between treatment and the repeated trials ( $F_{2,203} = 1.05$ , ns). To assess whether there was a more subtle effect of the treatment at the individual level, amount of improvement relative to pretest score was analyzed with analysis of covariance. No significant difference for treatment was found ( $F_{2,202} = 1.32$ , ns).

### DISCUSSION

Our results did not confirm the results of the unpublished study cited by Rauscher and Shaw (1998). Our results did not produce a Mozart effect, despite the use of the same Mozart performance used by Rauscher, *et al.* (1993, 1995) and the same measure used in Rauscher, *et al.* (1993), the Paper Folding and Cutting task from the Stanford-Binet Intelligence Measure. This result suggests that Rauscher and Shaw's explanation of why other studies have not produced a Mozart effect is incomplete. The results here also add to the list of nonconfirmations of Rauscher, *et al.* (1993, 1995).

Rauscher and Shaw (1998) cite the work of Rideout and colleagues (Rideout, Dougherty, & Wernert, 1998; Rideout & Laubach, 1996; Rideout & Taylor, 1997) as producing the Mozart effect. Steele, *et al.* (1997), however, has pointed out a procedural problem common to all the Rideout studies. The procedure in those experiments contrasts the effects of listening to

a Mozart sonata against listening to a progressive relaxation tape. The finding of a difference in performance between those two conditions could reflect a lowering of arousal in the progressive relaxation condition instead of an enhancement of cognitive ability by listening to the Mozart sonata. Olmstead, Shannon, Kirby, and Steele (1997) have obtained mood measures consistent with this suggestion in a procedure similar to that used by Rideout. Nantais (1997) has suggested also that mood or arousal mediates the production of differences between music and control groups. One suggestion for research is to include mood measures as a standard practice to examine the extent to which mood or arousal changes are operative in these experiments.

The recent study by Rideout, *et al.* (1998) illustrates a problem confronting investigators. Rideout, *et al.* added a second music condition, a performance by Yanni (*Acroyali/Standing in Motion*) and proposed that the work should produce an increase in spatial reasoning scores also, based on their judgement of musical similarity of the Mozart and Yanni selections. Rideout, *et al.* found that the improvement in spatial reasoning scores produced by listening to Yanni was greater than listening to Mozart, relative to the relaxation condition. Although Rideout, *et al.* had a rationale for their choice, it is not clear whether listening to Yanni should (or should not) produce an increase in spatial reasoning scores according to the Rauscher and Shaw hypothesis because the discoverers of the Mozart effect have not yet set public definitions for categorizing *specific* musical selections.

The results of Nantais (1997) and Wilson and Brown (1997) illustrate another design issue related to the lack of objective rules for classification of specific selections. Nantais reported that listening to music either by Mozart or by Schubert produced an increase in performance relative to a control condition of silence; however, the advantage of listening to Mozart disappeared when the control condition was a narrated story. Wilson and Brown (1997) reported that their musical control condition, which they described as "repetitive and lacking in complexity," produced improved performance, contrary to Rauscher, *et al.* (1995). The Mozart effect is a relative difference in performance, and both results suggest there must be a theoretical rationale for the choice of control group also.

The popular excitement about the Mozart effect rests on the inference that it is an easy home remedy to improve intellectual skills. Rauscher, Shaw, Levine, Ky, and Wright (1994) contributed to this excitement by linking the Mozart effect with the production of long-term cognitive enhancement through music education. The results reviewed here suggest that such a linkage is premature and that the basis of the Mozart effect is still unclear.

#### REFERENCES

- CARSTENS, C. B., HUSKINS, E., & HOUNSHELL, G. W. (1995) Listening to Mozart may not en-

- hance performance on the Revised Minnesota Paper Form Board test. *Psychological Reports*, 77, 111-114.
- KENEALY, P., & MONSEF, A. (1994) Music and IQ tests. *The Psychologist*, 7, 346.
- NANTAIS, K. M. (1997) Spatial-temporal skills and exposure to music: is there an effect, and if so, why? Unpublished master's thesis, Univer. of Windsor, Windsor, Ontario, Canada.
- NEWMAN, J., ROSENBAUGH, J. H., BURNS, K. L., LATIMER, B. C., MATOCHA, H. R., & VOGT, E. E. (1995) An experimental test of "the Mozart effect": does listening to his music improve spatial ability? *Perceptual and Motor Skills*, 81, 1379-1387.
- NGUYEN, N., SHAW, G. L., & TRAN, S. (1996) Exploring the Mozart effect: pilot work on the effect of a distractor task. Unpublished raw data, Univer. of California, Irvine.
- OLMSTEAD, C. G., SHANNON, R. A., KIRBY, J. L., JR., & STEELE, K. M. (1997) Does Mozart help or hinder relaxation? Poster presented at the 43rd annual meeting of the Southeastern Psychological Association, Atlanta, Georgia.
- RAUSCHER, F. H., & SHAW, G. L. (1998) Key components of the Mozart effect. *Perceptual and Motor Skills*, 86, 835-841.
- RAUSCHER, F. H., SHAW, G. L., & KY, K. N. (1993) Music and spatial task performance. *Nature*, 365, 611.
- RAUSCHER, F. H., SHAW, G. L., & KY, K. N. (1995) Listening to Mozart enhances spatial-temporal reasoning: towards a neurophysiological basis. *Neuroscience Letters*, 185, 44-47.
- RAUSCHER, F. H., SHAW, G. L., LEVINE, L. J., KY, K. N., & WRIGHT, E. L. (1994) Music and spatial task performance: a causal relationship. Presentation at the 102nd annual meeting of the American Psychological Association, Los Angeles, California.
- RAVEN, J. C. (1986) *Raven's Progressive Matrices*. San Antonio, TX: The Psychological Corp.
- RIDEOUT, B. E., DOUGHERTY, S., & WERNERT, L. (1998) Effect of music on spatial performance: a test of generality. *Perceptual and Motor Skills*, 86, 512-514.
- RIDEOUT, B. E., & LAUBACH, C. M. (1996) EEG correlates of enhanced spatial performance following exposure to music. *Perceptual and Motor Skills*, 82, 427-432.
- RIDEOUT, B. E., & TAYLOR, J. (1997) Enhanced spatial performance following 10 minutes exposure to music: a replication. *Perceptual and Motor Skills*, 85, 112-114.
- STEELE, K. M., BALL, T. N., & RUNK, R. (1997) Listening to Mozart does not enhance backwards digit span performance. *Perceptual and Motor Skills*, 84, 1179-1184.
- STEELE, K. M., BASS, K. E., & CROOK, M. D. (in press) The mystery of the Mozart effect: failure to replicate. *Psychological Science*.
- STOUGH, C., KERKIN, B., BATES, T., & MANGAN, G. (1994) Music and spatial IQ. *Personality and Individual Differences*, 17, 695.
- THORNDIKE, R. L., HAGEN, E. P., & SATTLER, J. M. (1986) *The Stanford-Binet Intelligence Scale: fourth edition*. Chicago, IL: The Riverside Publ. Co.
- WEEKS, S. P. (1996) The effect of music on abstract/visual reasoning performance in high school music and non-music students. (Doctoral dissertation, East Texas State Univer., 1995) UMI Dissertation Services No. 9600110.
- WILSON, T. L., & BROWN, T. L. (1997) Reexamination of the effect of Mozart's music on spatial-task performance. *The Journal of Psychology*, 131, 365-370.

Accepted April 5, 1999.