No Evidence for the Mozart Effect in Children

Rudi Črnčec, Sarah J. Wilson, and Margot Prior
School of Behavioural Science, University of Melbourne, Australia

The Mozart effect refers to claims that listening to Mozart-like music results in a small, short-lived improvement in spatiotemporal performance. Based on predominantly adult research that has shown equivocal findings, there has been speculation that the Mozart effect may have pedagogical benefits for children. The present study aimed to examine the Mozart effect in children and to evaluate two alternative models proposed to account for the effect, namely the trion model and the arousal-mood model. One hundred and thirty-six Grade 5 students (mean age \(11.0\) years) were exposed to three experimental listening conditions: Mozart piano sonata K. 448, popular music, and silence. Each condition was followed by a spatiotemporal task, and mood and music questionnaires. The results showed no evidence of a Mozart effect. Speculation about applications of the Mozart effect in children needs to be suspended until an effect can be reliably reproduced.

Received August 5, 2004, accepted April 26, 2005

The Mozart effect was originally taken to refer to predominantly adult reports of improved performance on tests of spatiotemporal function, occurring for a period of 10 to 15 min immediately after listening to the first movement of a Mozart piano sonata (K. 448) or similar compositions (Rauscher, Shaw, & Ky, 1993, 1995). Researchers have speculated that this effect may assist infants and children in developing their spatiotemporal and other cognitive abilities (Rauscher, 1999; Shaw, 2000). Particular emphasis has been placed on upper primary school, where topics putatively reliant on spatiotemporal abilities, such as proportional mathematics, are introduced (Shaw, 2000). Thus, evaluating the Mozart effect in children of this age is important in assessing possible pedagogical implications.

Use of the term Mozart effect has varied widely within the scientific literature (cf. Shaw, 2001). For example, in addition to the above definition, it has been used to describe the reported effects of (a) K. 448 on epileptiform activity in patients with epilepsy (Hughes, Daaboul, Fino, & Shaw, 1998), (b) K. 448 on maze learning in rodents (Rauscher, Robinson, & Jens, 1998), and (c) background music listening on spatiotemporal task performance (Ivanov & Geake, 2003). These diverse findings have not been shown to share a common underlying mechanism. Therefore, unifying them under a common term may create a false impression of the breadth and robustness of the Mozart effect. It is also important to differentiate the Mozart effect from a related line of research examining the effects of music lessons on cognitive abilities (cf. Hetland, 2000a). This study examined the Mozart effect as it was initially defined, that is, in relation to predictions about improved spatiotemporal performance following exposure to Mozart’s music.

There is considerable popular interest in the Mozart effect, much of which has resulted from the misreporting of scientific data and confusion about the parameters of the effect. For example, in the first study of the Mozart effect the results were reported as a short-term improvement in spatial IQ (Rauscher et al., 1993). This finding was subsequently described in the media as a long-term improvement in overall IQ, leading to the popular misconception that “Mozart makes you smarter.” Further complicating this picture, the term Mozart effect was trademarked by Campbell in 1997 as an inclusive term signifying proposed wide-ranging benefits of music in improving health, creativity, and intellectual abilities (Campbell, 1997). Campbell has written two books, neither of which are based on the scientific research outlined above (Campbell, 1997, 2000). A substantial industry based on this broader definition has since developed with compact discs, videos, and other items targeted especially at improving the abilities of infants and children. A thorough examination of research evidence for the Mozart effect in children will help determine the scientific basis of this industry.
Two main explanatory models of the Mozart effect have been proposed: initially the trion model (Leng & Shaw, 1991) and later the arousal-mood model (Thompson, Schellenberg, & Husain, 2001). The former is a mathematical model of neuronal firing in human cortex. It has been argued that as the trion model generates firing patterns with complex spatial and temporal patterns and Mozart’s music is structured according to complex, spatial and temporal patterns, exposure to Mozart primes the brain for spatiotemporal tasks (Shaw, 2000). Repetitive music is not thought to exert this effect owing to its “simple” structure (Rauscher & Shaw, 1998), although the terms “complex” and “simple” have not been musically operationalized within the research field. Schellenberg (2001) has argued that the assertion that K. 448 primes the brain for spatiotemporal performance is not supported by cognitive or neuropsychological research and is a “radical claim about cognitive processes” (p. 358). Given the relatively weak link between the trion model and the Mozart effect, studies have tended to be driven by empirical findings, rather than predictions from the trion model per se (e.g., Rauscher & Shaw, 1998).

The arousal-mood model claims that the Mozart effect is a function of the participant’s enjoyment of the stimulus and associated mild increases in arousal and positive mood. Husain, Thompson, and Schellenberg (2002) posited that any enjoyable stimulus may confer a small positive effect on spatiotemporal reasoning. Thus, they argued that the Mozart effect has nothing to do with music in general, or with Mozart per se. This model is supported by studies indicating that music can affect arousal and mood (e.g., Schmidt & Trainor, 2001; Sloboda & Juslin, 2001) and that arousal and mood can in turn influence cognition (e.g., Yerkes & Dodson, 1908; Isen, 2002). While this model reconciles the Mozart effect with well-reported scientific phenomena, a problem is raised. Arousal and mood are understood to affect performance on a range of cognitive tasks, but to date the Mozart effect has only been evident for spatiotemporal tasks. Several other cognitive abilities have been examined in relation to the Mozart effect, but none have shown significant improvements (cf. Chabris, 1999; Hetland, 2000b). Moreover, not all studies that have reported differences in mood or arousal following K. 448 have demonstrated a corresponding Mozart effect (e.g., Steele, Bass, & Crook, 1999; Steele, Dalla Bella, et al., 1999), raising concerns about the reliability of the Mozart effect.

The bulk of research examining the Mozart effect has been conducted with undergraduate students. Thus, speculation about benefits to children and infants has been largely based on the downward extension of adult findings. This is problematic for several reasons. First, it is generally accepted that findings considered to have applications for children should be directly evaluated within this population. Second, many extant adult studies, including the original reports by Rauscher et al. (1993, 1995), have methodological shortcomings (Fudin & Lembessis, 2004). These include poor control of potentially relevant variables such as arousal, mood, and music training and the use of “control” conditions that may actively depress spatiotemporal performance, such as relaxation instructions. Third, while the Mozart effect has been reproduced in adults by the original authors (e.g., Rauscher et al., 1995) and several others (cf. Chabris, 1999; Hetland, 2000b), there are approximately equal numbers of adult studies demonstrating or refuting the effect. In a meta-analysis of published adult research, Chabris (1999) found an overall nonsignificant effect of Mozart on spatiotemporal performance ($d = .14$). In a second meta-analysis, Hetland concluded that there was a medium-sized, short-lived Mozart effect ($d = .50$) (Hetland, 2000b). This second analysis included multiple unpublished manuscripts, many of which reported a large Mozart effect. None of these positive results has subsequently been published, raising concerns about Hetland’s summary data. Thus, in general the Mozart effect appears to be a small and unreliable effect in the adult population.

Only two previous studies have investigated the Mozart effect in children (Hallam, 2000; McKelvie & Low, 2002). In the first study, Hallam (2000) sought to replicate the Mozart effect in a naturalistic school environment. Participants were 8,120 children (age range, 10–11 years) recruited from 150 schools. Participants were tested in class groups, with the class teacher running the experiment from a script. A between-participants, posttest design was used where performance on two tests of spatial reasoning was assessed after exposure to either Mozart (K. 593), popular music, or a scientific discussion. The participants were tested simultaneously, with the experimental stimuli played on BBC radio. The analysis revealed no significant differences in spatial performance between the three conditions.

McKelvie and Low (2002) reported results from two experiments using designs that had previously shown a Mozart effect in adults. Study 1 employed a single-session,
between-participants, pre- versus posttest design with 55 participants (mean age = 12 years). Experimental stimuli were either K. 448 or a popular dance composition. McKelvie and Low (2002) included self-rated measures of enjoyment of the musical stimuli and previous music training. The results revealed no evidence of a Mozart effect, nor any interactions. In Experiment 2, K. 448 and popular music were contrasted with relaxation music. This study employed a between-participants, pre- versus posttest design, with 48 participants (mean age = 12.2 years). The results were similar to those of Study 1. McKelvie and Low (2002) utilized sex and music training as covariates in their data analyses, as both of these variables are thought to influence spatiotemporal abilities (cf. McKelvie & Low, 2002). Sex and music training were also employed as covariates in the present study.

Neither Hallam (2000) nor McKelvie and Low (2002) included a silence, or neutral control condition. Therefore, in both studies the null results may have been attributable to a general effect of listening to auditory material. Further, both studies utilized between-participants designs. Such a design makes it difficult to evaluate the predictions of the arousal-mood model, that is, the possible effect of individual responses to music on spatiotemporal function. In addition, between-participants designs do not control for individual differences in spatiotemporal ability across groups. Nonetheless, these studies suggest that the Mozart effect may not be evident in childhood populations.

In a related study, Ivanov and Geake (2003) assessed performance on a 10-item paper-folding task using a single-session posttest design with three background listening conditions: (a) K. 448, (b) Bach (DWV 916), and (c) silence. This use of continuous background music significantly changes the nature of the task, as attention is divided between the cognitive task and the music. Nevertheless, given the use of K. 448 and a spatiotemporal dependent variable, the results warrant consideration. Participants were 76 children with a mean age of 11.1 years.Musical experience was assessed by a self-report questionnaire before testing. Ivanov and Geake (2003) reported significantly better paper-folding performance for the Mozart and Bach conditions relative to the silence condition. No effect of music training was noted. The researchers speculated that (a) the results may be explained by the arousal-mood model, and (b) music may have brought cohesion to the background noise of the classroom, facilitating task performance (cf. Cash, El-Mallakh, Camberlain, Bratton, & Li, 1997). Ivanov and Geake (2003) could not evaluate whether the observed effects were related to the arousal-mood model, however, as they did not assess enjoyment, mood, or arousal following exposure to the stimuli.

**Rationale**

The present investigation sought to evaluate the Mozart effect in a pediatric population using a tightly controlled experimental paradigm to directly compare the two extant models of the effect. To our knowledge, such a comparison of models has not been previously conducted with either adults or children. We contrasted the effects of K. 448, popular music ("Zorba’s Dance"), and silence on spatiotemporal performance using a within-participants, repeated-measures design. "Zorba’s Dance" was chosen as the popular piece because this composition was predicted to be enjoyable and arousing to children and thus, according to the arousal-mood model, to improve spatiotemporal abilities. Moreover, as "Zorba’s Dance" was simple and repetitive, the trion model predicted no enhancement in spatiotemporal performance for this condition. K. 448 represented the composition predicted by the trion model to enhance spatiotemporal abilities. A silence condition was included to control for the possibility of a general listening effect. Spatiotemporal ability and self-reported preference, arousal, and mood were evaluated after exposure to each listening condition. Baseline visuospatial and musical abilities were assessed and each participant’s parents completed a questionnaire about their child’s music training. Baseline musical ability was also examined, as this variable is thought to be correlated with spatiotemporal functioning (e.g., Hassler, Birbaumer, & Feil, 1985, 1987) but has not been examined in studies of the Mozart effect to date. We predicted that any effects of music exposure on spatiotemporal ability would be attributable to (a) differences in baseline spatiotemporal ability, (b) preference for the listening condition, (c) mood following the listening condition, and/or (d) musical ability.

**Method**

**Participants**

One hundred and thirty-six Grade 5 children (56% male) were recruited to the study, with a mean age of 10.7 years (SD = .34; range, 9.5–11.6 years). Participants were recruited from six classes in three neighboring public schools from the same socioeconomic area in metropolitan Melbourne. Each school devoted a similar amount of time to group music lessons. Of the 136 children who participated in the study, there were 36 cases
of missing data. This was principally due to participants not attending one of the four testing sessions. In addition, data obtained from children with neurological, psychological, or learning disorders were excluded from the study. Missing cases were randomly distributed across schools and sex. The sample had sufficient experimental power to detect a small to medium-sized Mozart effect.

Stimuli and Measures

The K. 448 stimulus was the first movement of this sonata, lasting 8 min 23 s (Mozart, 1985). The popular music stimulus was created from a recording of “Zorba’s Dance” (LCD, 1998). “Zorba’s Dance” was modified using Cool Edit 2000 software (Johnston, 2000) to remove phrases from the composition that contained short verbalizations and to make the stimulus the same length as K. 448. In the silence condition, participants were instructed to sit quietly and were informed when 8 min 23 s had elapsed.

The primary measure of spatiotemporal reasoning was the Fitzgerald paper-folding test (Fitzgerald, 1978). This has been developed and normed in Australia for use with children in upper primary school and has been shown to demonstrate good internal consistency and test-retest reliability (Fitzgerald, 1978). It requires children to visualize paper being folded, punctured, and then unfolded and to select the correct answer from a multiple-choice array. It consists of a practice item and 20 test items administered in paper-and-pencil form and takes approximately 10 to 15 min to complete. The Fitzgerald paper-folding task satisfies Rauscher and Shaw’s (1998) criteria for a spatiotemporal task, in that it requires both spatial imagery and the temporal ordering of spatial components. It was not appropriate to use the paper folding and cutting subtest from the Stanford-Binet intelligence scale (4th edition), a test frequently used in adult Mozart effect research, as normative data are only available for children aged 12 and over (Thorndike, Hagen, & Sattler, 1986). Furthermore, given that the Mozart effect in adults is thought to wane after a period of 10 to 15 min (cf. Rauscher & Shaw, 1998), additional measures of spatiotemporal reasoning were not employed.

Children rated their preference for and familiarity with each experimental stimulus after it was played using 5-point Likert-type scales. These scales were created specifically for use in the present study and were presented in the form of a questionnaire.

Arousal and mood were measured using the 10-item affective reaction chart developed for children by Ainley (Ainley, Bretherton, & Sanson, 1994; Ainley, Hidi, & Tran, 1997; Ainley & Hidi, 2002). Participants were asked to rate the extent they experienced each of the 10 emotions after the listening condition using 5-point Likert-type scales. This measure derives its empirical base from Izard’s differential emotions theory (Izard, 1977) and his work on the measurement of emotions in children (cf. Izard, Dougherty, Bloxom, & Kotsch, 1974; Manstead, 1993). Particular emotions on the affective reaction chart, such as interest and surprise, provided a useful index of the participant’s level of arousal.

Other visuospatial measures administered during the pretest were the Porteous Mazes (Porteous, 1973), the Rey-Osterrieth complex figure test (Rey, 1941), and the Vandenberg three-dimensional mental rotation task (Vandenberg & Kuse, 1978). These were selected to provide a broader assessment of the participants’ visuospatial abilities. Musical ability was measured at pretest using the Bentley Measures of Musical Abilities (Bentley, 1966). This test assesses four areas: pitch discrimination, tonal memory, chord analysis, and rhythmic memory. These four scores can be combined to give a total musical ability score.

All parents and guardians were required to complete a questionnaire about their child’s music experience before the commencement of testing. This included questions about the presence and length of extracurricular music lessons to control for any effect of music training on the Mozart effect. Demographic information, handedness, and the presence of neurological, psychological, or learning disorders were also obtained.

Procedure

The relevant Human Research Ethics Committees approved this study, and formal consent was obtained before testing. Preexperimental testing took place in classroom groups with the children’s teacher present. The experimenter provided standard instructions for each test before its commencement, displayed on an overhead projector, so that the group moved through the tasks together. Talking was not permitted during test administration; however, due to the age of the participants, some chatter was allowed once everyone had finished a test. Teachers were encouraged to assist with disciplinary issues where required. Preexperimental testing duration was 1.5 hours.

Experimental testing commenced 1 week later with classes of children randomly allocated to one of six counterbalanced stimulus presentation conditions (see Table 1). The same procedures used in the preexperimental protocol were applied. Before the commencement of testing,
Participants were instructed in the use of the music and mood questionnaires and completed the Fitzgerald paper-folding test practice item. Following this, participants were informed that they would be exposed to a short period of music or silence. Participants were instructed not to talk, draw, or do anything else with their hands during this time; however, they were permitted to put their heads on their desks. The listening condition was preceded by a quick “stretch and wriggle” to assist the children to sit quietly. Participants were advised that immediately after presentation of the musical stimulus or silence they were to complete the Fitzgerald paper-folding test and questionnaires. The duration of experimental testing was approximately 40 min. This procedure was repeated three times at weekly intervals for each of the listening conditions. All musical stimuli were played on compact disc through a CDSONIC AE-240 Amplified Speaker System.

Results

Preexperimental Testing

Examination of the parent questionnaire indicated that 77 children (57.9%) were engaged in extracurricular music lessons. These children had been engaged in music lessons for an average of 32.7 months (SD = 22.40). Group mean scores for all pretest spatial and musical tasks were within average limits, indicating that the sample was representative in terms of spatial and musical abilities (see Table 2).

Experimental Testing

The Mozart effect. Data were normally distributed and suitable for parametric analyses. To examine whether there was a significant change in paper-folding scores as a function of music listening condition (Mozart, popular music, or silence), a within-participants repeated-measures ANCOVA was performed. Sex and the presence of music training were included as covariates in this analysis. The main effect of listening condition on paper folding was not significant, $F(2, 188) = 1.6$, $p = .21$. The main effects of sex and music training were also not significant, $F(1, 94) = .10$, $p = .76$, and $F(1, 94) = .43$, $p = .51$, respectively. Neither sex, $F(2, 188) = 1.42$, $p > .05$, nor music training, $F(2, 188) = .87$, $p > .05$, significantly adjusted paper-folding scores. These results are shown in Figure 1.

As the Fitzgerald paper-folding test was administered on four separate testing sessions (Pretest, Posttest 1, Posttest 2, and Posttest 3), a mixed repeated-measures ANOVA was performed to examine the effect of practice, with counterbalanced group as the between-participants factor. There were six levels of counterbalanced group (see Table 1). The main effect of counterbalanced group was not significant, $F(5, 86) = .99$, $p > .05$, nor was the testing session by counterbalanced group interaction, $F(15, 258) = 1.47$, $p > .05$. However, a significant main effect of testing session was observed, $F(3, 258) = 13.25$, $p < .0001$. Post hoc simple contrasts indicated that paper-folding performance differed significantly between Pretest and Posttest 3, $F(1, 86) = 19.29$, $p < .0001$, and between Posttest 1 and Posttest 3, $F(1, 86) = 14.03$, $p < .0001$, but not between Posttest 2 and Posttest 3, $F(1, 86) = 2.62$, $p > .05$.

The possibility that participant improvement across testing sessions concealed a Mozart effect was further evaluated by examining the effect of listening condition (K. 448, popular, and silence) on Posttest 1 experimental scores. These scores are least likely to be affected by practice. This effectively changed the design to a between-participants posttest comparison, as commonly used in previous research. The results revealed no effect of listening condition on paper-folding performance during Posttest 1, $F(2, 118) = 1.04$, $p = .36$, supporting the absence of a Mozart effect.

Preference, mood, and arousal effects. A further possibility for the lack of a Mozart effect was that the music

![FIG. 1. Mean paper-folding scores following exposure to each of the experimental listening conditions.](image-url)
was not associated with increased preference, positive mood, or arousal, as measured using the 5-point rating scales. To explore differences in the children’s subjective responses to K. 448, popular music, and silence, a repeated-measures within-participants ANOVA was performed across the three listening conditions for each rating category. Simple post hoc contrasts were used to explore significant effects of listening condition. As shown in Table 3, ratings of preference, $F(2, 204) = 100.33, p < .0001$; happiness, $F(2, 198) = 61.93, p < .0001$; interest, $F(2, 202) = 87.74, p < .0001$; and surprise, $F(2, 206) = 18.46, p < .0001$, were significantly higher following popular music compared to K. 448 or silence. This supports the notion that the popular condition served its intended purpose of inducing positive mood and arousal; yet, despite this, it did not confer an advantage on paper-folding performance. Participants reported that K. 448 was significantly more complex than popular music or silence, $F(2, 204) = 47.64, p < .0001$. Thus, K. 448 served as a good measure of the trion model in children.

Participants also reported that K. 448 was less familiar than either popular music or silence, $F(2, 204) = 71.19, p < .0001$. Given that familiarity of the experimental stimulus has not been previously investigated in Mozart effect studies but may be important, this variable was included in subsequent analyses.

**Predictors of spatiotemporal performance.** To examine the contribution of specific within-participants variables to spatiotemporal ability, a multiple regression
analysis was performed for each listening condition. In all analyses, the dependent variable was paper-folding score following exposure to the experimental stimulus. Independent variables were selected on the basis of a possible association with paper-folding score and included (a) self-reported happiness (positive mood), (b) preference for the listening condition, (c) familiarity of the stimulus, (d) musical ability, and (e) pretest paper-folding performance. This last variable represented an important control for individual differences in baseline spatiotemporal ability in this study. Pearson $r$ correlations between these variables are shown in Table 4.

As indicated in Table 5, the results of the K. 448 and popular music multiple regression analyses were similar. That is, $R$ was significantly different from zero,

### Table 4. Correlations between variables used in the K. 448, popular music, and silence multiple regression analyses.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Condition</th>
<th>Pretest paper-folding score</th>
<th>Happiness</th>
<th>Preference</th>
<th>Familiarity</th>
<th>Musical ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K. 448</td>
<td>.63**</td>
<td>.04</td>
<td>.09</td>
<td>-.06</td>
<td>.29*</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td>.65**</td>
<td>.05</td>
<td>.09</td>
<td>-.02</td>
<td>.26*</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.66**</td>
<td>-.09</td>
<td>-.05</td>
<td>-.00</td>
<td>.28*</td>
</tr>
<tr>
<td>Posttest paper-</td>
<td>K. 448</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>folding score*</td>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest paper-</td>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>folding score</td>
<td>Silence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Happiness</td>
<td>K. 448</td>
<td>.59**</td>
<td>.02</td>
<td>.05</td>
<td>.14</td>
<td>.27*</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference</td>
<td>K. 448</td>
<td>.38**</td>
<td>.02</td>
<td>.05</td>
<td>-.07</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiarity</td>
<td>K. 448</td>
<td>.15</td>
<td>.01</td>
<td>.05</td>
<td>.27*</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Musical ability</td>
<td>K. 448</td>
<td>.14</td>
<td>.01</td>
<td>.05</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest paper-</td>
<td>K. 448</td>
<td>.65**</td>
<td>.08</td>
<td>.09</td>
<td>.33</td>
<td>.33</td>
</tr>
<tr>
<td>folding score</td>
<td>Popular</td>
<td>.66**</td>
<td>.13</td>
<td>.09</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.62**</td>
<td>.62</td>
<td>.09</td>
<td>.35</td>
<td>.35</td>
</tr>
</tbody>
</table>

*Refers to posttest score collapsed across all counterbalanced presentations of each listening condition.

*p < .01. **p < .0001.

### Table 5. Multiple regression analyses of within-participants variables on paper-folding scores following exposure to K. 448, popular music, or silence.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Condition</th>
<th>$B$</th>
<th>$\beta$</th>
<th>$sr,^2$ (unique)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happiness</td>
<td>K. 448</td>
<td>.27</td>
<td>.08</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td>.31</td>
<td>.10</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.11</td>
<td>.04</td>
<td>.00</td>
</tr>
<tr>
<td>Preference</td>
<td>K. 448</td>
<td>.25</td>
<td>.07</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td>-.01</td>
<td>-.00</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.04</td>
<td>.01</td>
<td>.00</td>
</tr>
<tr>
<td>Familiarity</td>
<td>K. 448</td>
<td>-.36</td>
<td>-.09</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td>-.14</td>
<td>-.05</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.01</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>Musical ability</td>
<td>K. 448</td>
<td>.07</td>
<td>.13</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>Popular</td>
<td>.05</td>
<td>.09</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.08*</td>
<td>.15</td>
<td>.02</td>
</tr>
<tr>
<td>Pretest paper-</td>
<td>K. 448</td>
<td>.65**</td>
<td>.60</td>
<td>.33</td>
</tr>
<tr>
<td>folding score</td>
<td>Popular</td>
<td>.66**</td>
<td>.63</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>Silence</td>
<td>.62**</td>
<td>.62</td>
<td>.35</td>
</tr>
</tbody>
</table>

Note. K. 448 regression: $R^2 = .44$, adjusted $R^2 = .41$, $R = .66*$; popular music regression: $R^2 = .44$, adjusted $R^2 = .41$, $R = .66*$; silence regression: $R^2 = .46$, adjusted $R^2 = .44$, $R = .68**$.

*p < .05. **p < .0001.
Mozart, popular music, or silence. Predictions made by the trion model were not upheld. Despite being rated as more complex by the children, K. 448 did not enhance spatiotemporal performance compared to repetitive music or silence. Taken together, theoretical concerns with the trion model (cf. Schellenberg, 2001) and the lack of behavioral data demonstrating a Mozart effect in children (Hallam, 2000; McKelvie & Low, 2002) suggest that complex music does not prime children’s brains for spatiotemporal tasks. In contrast, predictions of the arousal-mood model were partly upheld. Exposure to popular music was associated with enhanced positive mood, arousal, and increased preference; however, these changes did not result in improved spatiotemporal performance. Rather, participants in this study showed short-term stability in their performance of the spatiotemporal task, with pretest spatiotemporal performance most strongly predicting posttest experimental scores.

Findings from this study are consistent with and extend those of McKelvie and Low (2002) and Hallam (2000). Together these investigations provide corroborative evidence that the Mozart effect does not exist in childhood populations and is not a function of subjective responses to music or the effects of prior music training. The results of the present study are at odds with those of Ivanov and Geake (2003), who demonstrated improved paper-folding performance in upper-primary-school-age children after continuous listening to K. 448 or the music of Bach. Although Ivanov and Geake’s results need to be replicated before they can be considered conclusive, it is interesting to consider the present findings in light of these data. Specifically, the possibility emerges that while a pretest musical stimulus is insufficient to produce a small enhancement in spatiotemporal reasoning in children, continuous exposure to music may produce an effect. This will be discussed further below.

### TABLE 6. Multiple regression analysis of Bentley measures of musical ability subtest scores on posttest paper-folding scores following exposure to silence, and correlations between variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Paper-folding score</th>
<th>Pitch</th>
<th>Tunes</th>
<th>Rhythm</th>
<th>B</th>
<th>β</th>
<th>$\text{sr}^2$ (unique)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tunes</td>
<td>.23*</td>
<td>.36**</td>
<td></td>
<td></td>
<td>.08</td>
<td>.07</td>
<td>.00</td>
</tr>
<tr>
<td>Rhythm</td>
<td>.25**</td>
<td>.28*</td>
<td>.23*</td>
<td></td>
<td>.39*</td>
<td>.20</td>
<td>.04</td>
</tr>
<tr>
<td>Chords</td>
<td>.16</td>
<td>.12</td>
<td>.17</td>
<td></td>
<td>.06</td>
<td>.16</td>
<td>.02</td>
</tr>
</tbody>
</table>

*Note. $R^2 = .13$, adjusted $R^2 = .09$, $R = .35*$.*

*p ≤ .05, **p < .001.

$F(5, 98) = 15.10$, $p < .0001$, and $F(5, 100) = 15.39$, $p = .0001$, respectively, with pretest paper-folding score the only variable to significantly contribute to the prediction of posttest paper-folding score ($sr^2 = .33$ and .35, respectively). In these analyses, scores on the five independent variables predicted 40.6% and 40.7%, respectively, of the adjusted variability in spatiotemporal performance.

$R$ was significantly different from zero in the silence multiple regression analysis, $F(5, 97) = 16.83$, $p < .0001$, with pretest paper-folding performance again strongly predicting posttest paper-folding performance ($sr^2 = .35$). In this analysis, musical ability also significantly contributed to the prediction of paper-folding score ($sr^2 = .02$). The five independent variables accounted for 43.7% of the adjusted variability in spatiotemporal performance (see Table 5). In order to evaluate which aspects of music ability were contributing to paper-folding performance, a further multiple regression analysis was conducted on posttest paper-folding scores following exposure to silence. The independent variables were the four subtest scores of the Bentley Measures of Musical Ability, namely, pitch, tunes, chords, and rhythm. As summarized in Table 6, $R$ was significantly different from zero, $F(4, 102) = 3.65$, $p = .01$. Only the rhythm subtest contributed significantly to the prediction of paper-folding scores ($sr^2 = .04$). Altogether, these four independent variables predicted 9.1% of the adjusted variability in paper-folding scores.

### Discussion

The results of this study indicated no evidence of a Mozart effect in upper-primary-school-age children. Children performed no differently on tests of spatiotemporal reasoning following passive exposure to Mozart, popular music, or silence. Predictions made by the trion model were not upheld. Despite being rated as more complex by the children, K. 448 did not enhance spatiotemporal performance compared to repetitive music or silence. Taken together, theoretical concerns with the trion model (cf. Schellenberg, 2001) and the lack of behavioral data demonstrating a Mozart effect in children (Hallam, 2000; McKelvie & Low, 2002) suggest that complex music does not prime children’s brains for spatiotemporal tasks. In contrast, predictions of the arousal-mood model were partly upheld. Exposure to popular music was associated with enhanced positive mood, arousal, and increased preference; however, these changes did not result in improved spatiotemporal performance. Rather, participants in this study showed short-term stability in their performance of the spatiotemporal task, with pretest spatiotemporal performance most strongly predicting posttest experimental scores.

Findings from this study are consistent with and extend those of McKelvie and Low (2002) and Hallam (2000). Together these investigations provide corroborative evidence that the Mozart effect does not exist in childhood populations and is not a function of subjective responses to music or the effects of prior music training. The results of the present study are at odds with those of Ivanov and Geake (2003), who demonstrated improved paper-folding performance in upper-primary-school-age children after continuous listening to K. 448 or the music of Bach. Although Ivanov and Geake’s results need to be replicated before they can be considered conclusive, it is interesting to consider the present findings in light of these data. Specifically, the possibility emerges that while a pretest musical stimulus is insufficient to produce a small enhancement in spatiotemporal reasoning in children, continuous exposure to music may produce an effect. This will be discussed further below.
In this study, musical ability was found to contribute significantly to the prediction of paper-folding performance in the silence condition. While the contribution was small \((r^2 = .02)\), this suggests that children with greater musical ability also have greater spatiotemporal ability. Indeed, the Bentley Measures of Musical Ability total score was positively correlated with several other measures of visuospatial skills taken during the pretest (see Table 2). A link between musical and visuospatial abilities in children has been previously well demonstrated (Barret & Barker, 1973; Manturzewska, 1978; Karma, 1979; Hassler et al., 1985, 1987; Lynn, Wilson, & Gault, 1989; Nelson & Barressi, 1989; Gromko & Poorman, 1998); however, further research is required to elucidate specific factors that underpin this association. The relationship between musical and visuospatial skills highlights the importance of controlling for musical ability in studies of the Mozart effect and raises the possibility that previous research reporting a Mozart effect may be confounded by this variable.

The present results indicated that rhythmic ability was the musical skill most strongly associated with spatiotemporal performance. This result has not been previously described, likely reflecting the absence of rhythmic measures in research investigating associations between musical and visuospatial abilities. There is, however, some evidence linking rhythmic ability with other nonmusical abilities. For example, Lynn et al. (1989) reported that children's performance on both rhythm and pitch discrimination tasks was positively associated with measures of general intelligence. Rhythm discrimination has also been associated with reading and spelling abilities (Douglas & Willatts, 1994). There are at least two possible explanations for the association between rhythmic and spatiotemporal abilities observed in this study. First, both the rhythm discrimination task from the Bentley Measures of Musical Ability and the Fitzgerald paper-folding task rely on working memory function. This notion is supported by the finding that the tonal memory task from the Bentley Measures of Musical Ability also correlated with posttest paper-folding performance and has a predominant working memory component (see Table 6). In other words, all of these tasks require the stimulus to be “held in mind” over a period of several seconds while a discrimination judgment is made. Alternatively, both the rhythm discrimination and paper-folding tasks may have a visuomotor component, suggesting that general motor skills may partly account for the observed association. The links between rhythmic and visuospatial abilities offer exciting avenues for future research.

### General Discussion

#### Methodological Considerations in the Current Design

The null results of this study may reflect methodological issues. First, the use of a single measure of spatiotemporal reasoning represents a methodological concern common to Mozart effect studies. Specifically, the use of additional spatiotemporal measures could improve the validity of results, excepting that the short-lived nature of the Mozart effect makes this unfeasible unless multiple experiments are conducted. Second, the readministration of the Fitzgerald paper-folding test on several occasions led to a small practice effect across testing sessions, which may have “washed out” any Mozart effect. We consider this unlikely for several reasons. The experimental conditions were presented in counterbalanced order, and thus, the effects of practice were equivalent across conditions. Moreover, the nonsignificant interaction between testing session and counterbalanced group suggests that the observed practice effect did not mask an effect of listening condition. Examination of the results did not indicate that the participants reached a ceiling level of performance on the paper-folding task. Therefore, any improvements associated with exposure to K. 448 or popular music should have been detectable above those associated with practice of the task. In addition, post hoc analysis of the first experimental session scores in isolation did not demonstrate evidence of a Mozart effect. This suggests that the broader results of the study are robust.

#### Developmental Considerations

Developmental factors could explain why children's performance on spatiotemporal tasks may not be influenced by small fluctuations in arousal and mood brought about by pretest exposure to music. The major explanatory models of the Mozart effect may be invalid in child populations. This, in turn, raises the dual possibility that either the Mozart effect is exclusively an adult effect or the Mozart effect does not exist at all. In the case of the former, children's performance on spatiotemporal tasks may not be influenced by small fluctuations in arousal and mood. The present study supports the proposition of McKelvie and Low (2002) that children exhibit short-term stability in their performance on these tasks. Further, while the general principles of the Yerkes-Dodson law of arousal (1908) are thought to apply to developmental populations, to date few studies have directly examined the role of arousal and mood in mediating children's cognitive...
performance. After evaluating the effects of mood on children's memory and impression formation, Forgas, Burnham, and Trimboli (1988) speculated that “There are likely to be profound, and as yet not fully explored differences between adults and children in the way mood states influence their cognitive abilities” (p. 703). Theories of adult responses to arousal and mood may not be directly applicable to children. For example, researchers have noted that background music aimed at reducing levels of arousal may exert beneficial effects on children's schoolwork, while arousing music can disrupt performance (Hallam, Price, & Katsarou, 2002).

There are developmental differences between the brains of children and adults that may account for different responses to positive mood and arousal. Generally speaking, the brains of children aged between 9 and 12 years are undergoing several neuronal developmental processes, including a decrease in synaptic density, axonal elimination, changes in global cerebral metabolism, and myelination of axonal fibers in the telencephalon (Brown, Keynes, & Lumsden, 2001). Of particular relevance is the lack of maturation of cortical regions thought to be important in spatiotemporal reasoning, including the parietal and frontal lobes (Klingberg, Forssberg, & Westerberg, 2002) and connections between these regions (Lambe, Krimer, & Goldman-Rakic, 2000; Klingberg et al., 2002). It is understood that dopaminergic projections to the frontal cortex continue to develop during adolescence and into early adulthood (Benes, 2003). Increased dopaminergic activity in frontal regions associated with positive mood may partly account for the Mozart effect in adults (Isen, 2002). Dopaminergic projections to the frontal cortex are underdeveloped in children, potentially reducing this effect.

Alternatively, it is conceivable that music may need to be present continually in order to exert a small effect on spatiotemporal performance in children (Ivanov & Geake, 2003). The affect regulation and attentional systems of children are functionally immature (Eisenberg & Fabes, 1999; Manly et al., 2001); thus mood states induced by music may rapidly subside once the music has ceased. There is little research on the durability of experimentally induced moods in pediatric populations; however, there is some indication that in adults, mood states induced in the laboratory may dissipate after approximately 15 min (for review see Brenner, 2000). Existing research examining background classroom music has not demonstrated reliable effects on cognitive or academic performance (Črnčec, Wilson, & Prior, in press). The present study is unable to shed light on these issues, and further research with adults and children is clearly warranted.

Even if we assume that the arousal-mood model can accurately account for the Mozart effect, compositions by Mozart may not be the best music to play to children to enhance mood and arousal. Results from the present study indicate that children enjoyed Mozart to the same extent as sitting in silence for 8 min. Further, it is conceivable that playing Mozart or, more generally, classical music to children may impair cognitive performance, as it may induce negative mood (cf. O’Hanlon, 1981). The observation that children and adolescents do not, by and large, enjoy classical music is well documented (e.g., Hargreaves, Comber, & Colley, 1995; LeBlanc, Sims, Siivola, & Obert, 1996; Hargreaves & North, 1997). The contemporary success of piped classical music as a deterrent for the congregation of children and adolescents in public places is testimony to this (Grabosky, 1995; Hughes, McLaughlin, & Muncie, 2002).

Given the lack of a Mozart effect in children, and the equivocal nature of the effect in adults, at present there would appear to be no immediate spatiotemporal benefits of exposing children to Mozart. Simplistic solutions like the Mozart effect can create false impressions about child development and mislead well-meaning parents into purchasing products they might not have otherwise. Moreover, given promising research investigating the importance of musical interactions between caregiver and infant (Trehub, 2002; Trevarthen & Malloch, 2002), and the sheer joy that music can bring to one’s life, it is important that a null Mozart effect does not overshadow the other broader benefits of music.

Author Note

We thank James Canty, Mel Gallagher, and Ben Elias for their help in conducting this study. We also acknowledge all participants and teachers for generously giving of their time. The helpful comments provided by three anonymous reviewers on an earlier draft of this article were also appreciated. This article is based on research Rudi Črnčec undertook toward his Doctor of Psychology (Clinical—Child Specialisation) degree at the University of Melbourne.

Address correspondence to: Rudi Črnčec, MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, South Penrith DC, NSW 1797, Australia. E-MAIL r.crncc@uws.edu.au
References


