Creative Thinking: Music Improvisational Skills Development among Elementary School Students

Mark T. Kiehn, Assistant Professor, University of Wisconsin-Green Bay, kiehnm@uwgb.edu

Abstract

The purpose of this study was to investigate the music improvisational skill of students in grades 2, 4, and 6. Randomly selected participants from 3 elementary schools (N=60) were given one measure, the Vaughan Test of Musical Creativity; and two independent judges scored student responses. Rated dimensions were originality, rhythmic interest, and melodic interest. There were no significant gender differences between males and females on test scores; however, a significant grade level difference emerged with grade 2 students scoring significantly lower than grade 4, and slightly lower than grade 6 students. The results indicate a creative thinking growth stage may exist from grade 2 to 4, followed by a developmental leveling (no significant change in test scores) between grade 4 and grade 6.

Introduction

Improvisation and composition activities are established components of the school music education curriculum (Consortium of National Arts Education Associations, 1994; Hoffer, 2005; Kratus, 1996; Music Educators National Conference, 1996; Shehan-Campbell & Scott-Kassner, 2006). Although recent survey data reveal that music teachers report incorporating all nine National Standards in the music classroom, little instructional time was actually devoted to implementing creative thinking activities on a regular basis (Bell, 2003; Brophy, 2002b; Orman, 2002; Strand, 2006). Results of these studies indicate that teachers may feel inadequately prepared or uncomfortable teaching creative thinking activities.

The creative thinking process in music appears to be driven by a product intention (goal or vision of the individual) observable in music improvisation, composition, and analysis activities (Webster, 1987, 1988). Webster (1990) developed a theory of creative thinking in music, and at the heart is a unique type of thinking termed divergent thinking. Divergent thought is rooted in finding many possible answers to a particular problem or open-ended task. According to this theory, divergent thinking dimensions include originality (uniqueness of the musical ideas), musical extensiveness (number of ideas generated or fluency), and flexibility (ease of shifting within parameters such as loud/soft dynamic levels or high/low in pitch).

With regard to the musical extensiveness dimension or musical fluency, for example, creative children may generate a number of ideas when asked to create a sound story with their voice or other instrument. During this particular open-ended improvisation activity, ideas interact with other elemental concepts such as the student’s ability to create melodies, rhythms, or develop a phrase through the use of various dynamic and pitch levels. Further examination of the literature reveals several tests of divergent thinking ability in both music and artistic drawing based upon an open-ended task format and dimensions of originality, fluency, or flexibility (Gorder, 1976; Torrance, 1974, 1998; Vaughan, 1971; Vold, 1988; Webster 1977, 1983, 1987, 1994). Comparisons of test scoring criteria show authors generally agree on definitions for originality as the uniqueness of response and fluency as the sheer number of responses but not on the flexibility dimension.
Review of Related Literature

Although there is an expanding body of recent studies on students’ creative thinking through both music composition and improvisation activities, the music education literature reveals few investigations spanning five or more years (Brophy, 1999; Kiehn, 2000; Kratus, 1985, 1989; Swanwick and Tillman, 1986). The relative of recent research on improvisational skills development makes further study important, since both Kratus and Swanwick and Tillman examined creative thinking through music composition, not improvisation activities. In all, the results from the literature vary, but reveal that changes in the musical characteristics of children’s compositions and improvisations may be related to age.

Music compositional skill level was found to increase with age in an early study by Kratus (1985). Kratus explored the elemental characteristics (rhythm, melody, motive, phrase) of music compositions produced by 80 students age 5 through 13. Each subject was provided 10 minutes to compose and rehearse a short piece on a hand-held electronic keyboard. He reports that a student’s understanding of the rhythmic organization, melodic organization, and phrase construction of music increases consistently from age 5 through 11 (grades K-6 approximately). Kratus further notes that rhythmic development was very rare for all age groups, and students developed melodic concepts faster than rhythmic concepts. Songs composed by 13-year-olds in the study was rated slightly or significantly lower on almost all characteristics than those of the 11-year-olds; which may indicate a plateau in the acquisition of musical syntax during that period.

Kratus (1989) investigated the amount of time that age 7 through 11 students spent on various compositional processes while creating a melody. The author provided 60 students 10 minutes to compose a song on an electronic keyboard, and asked each participant to play her or his song and repeat it. Analyses focused on the amount of exploration, development, repetition, and silence students used while composing. His results suggest that composition activity among 7-year-olds was very similar to the act of improvisation, using time for exploration rather than modifying ideas. Kratus further states repetition is a necessary process; and 9- and 11-year-olds are capable of using exploration, development, and repetition in a manner similar to those used by adult composers.

A study of the undirected creative musical processes of children aged 3 to 11 was conducted by Swanwick and Tillman (1986). Each participant (N=48) was given a variety of open-ended musical tasks; and phrases were recorded nine times each year, yielding 745 compositions over four years. The student compositions were rated on four comprehensive domains of musical criticism: materials, expression, form, and value. The results showed there was a significant positive correlation between age and the judges’ ratings. The authors developed a theoretical model showing that domains of musical criticism characterize eight levels of musical development from Sensory level one (experimentation/exploration) to Systematic level eight (meta-cognitive mode) possible among students 15 years of age and older.

Age-related differences in the musical improvisations of children age 6 through 12 were examined by Brophy (1999). Participants (N=280) improvised three 8-measure melodies using the C-pentatonic scale on an Orff alto xylophone as the B, C, and D sections of a seven-part rondo (ABACADA form). Students were rated on melody, rhythm, and structure dimensions; and the results revealed a significant age difference for improvisation scores. Brophy concluded that a dynamic stage of development might exist from ages 6 through 9 (with a significant change between ages 8 and 9), followed by a developmental plateau from ages 9 through 11, with improvisational skills development apparently resuming at age 12. Brophy suggests the data indicates a developmental trend towards greater use of repeated and developed motives,
increased rhythmic conformity, and greater employment of organizing structural feature. He also reports the effect of gender was not significant.

Kiehn (2000) compared both music improvisational skills and creative thinking scores of 89 students in grades 2, 4, and 6. All participants completed the Vaughan Test of Musical Creativity (containing six open-ended improvisation tasks), and the Torrance Tests of Creative Thinking, Figural Form A (containing three open-ended artistic drawing tasks). The student responses on both measures where rated on originality and fluency dimensions. Kiehn reported grade level differences for both music improvisation scores and creative thinking scores with grade two students scoring significantly lower than grade 4 and grade 6 students. He also reported significant gender differences for music improvisation scores, with boys scoring higher than girls. Kiehn’s findings support Brophy’s (1999) proposition of a dynamic growth stage in improvisational skill from grade 2 to 4, followed by a leveling-off period (no significant change in test scores) between grade 4 and grade 6.

A recent longitudinal study conducted by Brophy (2005) focused on the development of a group of 62 students’ melodic improvisations over the course of three years. Participants of the study improvised, as part of a class rondo activity for Orff instruments, three 8-measure melodies each year. A total of 558 improvisations were collected and examined for musical characteristics of melody, rhythm, and structure. Brophy found significant age effects for repeated melodic and rhythmic motives, sense of pulse, and antecedent/consequent phrases. In sum, he reports the participants’ improvisations at ages 8 and 9 sounded more musically organized than at age 7. His hypothesis of a dynamic stage of music improvisational skill development between ages 6 and 9 (Brophy 1999, 2002a) was supported by the results of the 2005 study.

The results of music education research remain inconclusive when authors examined the relationship between creative thinking and gender (Auh, 1995; Baltzer, 1990; Brophy, 1999; Kiehn, 2000; Schmidt and Sinor, 1986; Swanner, 1985; Webster, 1977), which may suggest further investigation. Gender differences were significantly related to creative thinking test scores in four studies (Baltzer, 1990; Kiehn, 2000; Schmidt & Sinor, 1986; Webster, 1977); and researchers have posed various explanations such as differences in sampling procedures, types of measures used, testing procedures, and scoring procedures. Common to Baltzer (1999), Kiehn (2000), Schmidt and Sinor (1986), and Webster (1977) were the test dimensions used; and the originality and fluency subscale test scores in each of these studies were significantly higher for males. Gender differences may or may not emerge if scoring procedures include specifically defined musical characteristics or elements, and are not based entirely on generally defined dimensions.

The present study continues the line of research designed to further investigate trends in music improvisational skill development reported in the literature. The grade level range in music creativity research is often limited (1 to 3 years), and only a few researchers (Brophy, 1999; Kiehn, 2000) have investigated improvisation development spanning 5 or more years. Since Brophy and Kiehn suggest changes in improvisational skills may occur as children age; further replication research examining improvisation skills development appears appropriate. The primary purpose of this study was to compare the music improvisational skills of students in grades two, four, and six. The following research questions were addressed:

1. Can music improvisational skills be reliably measured?

2. Are there significant grade level differences for scores on a test of music improvisational skills?

3. Are there significant gender differences for scores on a test of music improvisational skills?
Methodology

Participants

Three schools were randomly selected from a total of 34 elementary schools within two large public school districts in Eastern Wisconsin. The selected schools represented suburban and urban geographic locations. Study participants were then selected from a total of 725 second, fourth, and sixth grade students representing the three randomly selected schools. Within each school, 24 students were selected (4 girls and 4 boys each from grades 2, 4, and 6) resulting in an initial sample size of 72 students. Individuals who had prior private training in music were eliminated from the final population pool.

Measure

All students receive general music instruction during school hours for 60 minutes each week from the school music specialist. The standards-based music curriculum in each school is comprehensive, and includes introductory improvisation experiences during each grade level on Orff melodic and percussion instruments. One measure was administered by the researcher to all participants (N=60) individually during their regular music class time. Each participant completed the Vaughan Test of Musical Creativity (1971), a measure of music improvisational skill. The participants were tested individually for 20 minutes in a quiet room, and responses to the Vaughan Test of Musical Creativity (TMC) were audio tape recorded for later scoring.

The TMC is comprised of six open-ended music improvisation tasks, and was selected as the criterion measure primarily because the test tasks align with activities found in the music education curriculum for both primary and intermediate grades (Consortium of National Arts Education Associations, 1994; Mead, 1994; Steen, 1992; Warner, 1991; Wisconsin Department of Public Instruction, 1989, 1997). The original test appears quite short in length (8 minutes of total administration time) and therefore somewhat limited in the data it creates. Kiehn (2000) lengthened the tasks and extended the total administration time to 10 minutes, which apparently enabled subjects to create a multiplicity of different responses.

Minor modifications to the test content were made, and response time extended in the present study, resulting in a total administration time of 20 minutes. The first and second test tasks served as warm-up activities and were the only items not scored. This helped to establish a relaxed and comfortable test climate for the participant, and the tester made it clear to all participants that they were “playing a musical game” throughout each session, to help reduce the possibility of test anxiety. The first warm-up task introduces the student to playing a steady beat on the drum while the tester plays a duple rhythm pattern on the claves. In Task 2, the tester performs a steady beat on the claves and asks the student to improvise a rhythm on the drum.

Tasks three through six are scored. The third task requires the student to create an answer rhythm on the drum (consequent phrase) to six different four-measure phrase questions in triple meter performed on claves by the tester. In Task 4, the tester plays several eight-beat melodic phrases on the bells (F-sharp pentatonic scale), and asks the student to respond to each melody with an answering tune on the black bells only. The tester plays a duple meter C-G broken bourdon on bells in Task 5, and asks the student to “improvise or make up a melody” on an Orff alto xylophone. The final task requires the participant to make up a piece showing how she or he feels during a thunderstorm, using their voice or any combination of the other instruments used during the testing session.
The researcher-modified scoring procedures yield an originality subscore, rhythmic interest subscore, melodic interest subscore, and a composite score (sum of the three subscores) for a composite score range of 0-45. The originality dimension is judged through a continuous five-point rating scale for tasks 3-6, rhythmic interest is judged through the same rating scale for tasks 3-5, and melodic interest for tasks 4-5. Similar to other studies in the music education literature, the originality dimension was defined as the uniqueness or novelty of response (Gorder, 1976; Webster, 1987, 1990, 1994); and definitions key to judging musical characteristics were used for the rhythmic interest and melodic interest dimensions (Brophy, 2005; Priest, 2001). The rhythmic interest subscore was specifically based on: (a) pulse adherence, (b) developed antecedent/consequent phrases, and (c) the presence of developed rhythmic motives. Melodic interest subscores were based on: (a) the presence of phrases, (b) developed antecedent/consequent phrases, and (c) the presence of developed melodic motives.

**Procedure**

The entire study, including test administration, scoring procedures, and adjudicator training sessions encompassed approximately 20 weeks. Two independent judges were trained to score the TMC during a pilot study prior to the main study. To enhance reliability, the test was first administered to nine students who were not included in the main study. Student responses in the pilot study were tape recorded for later scoring. Each judge listened to the products in a different random order on the audiotape, and each judge considered the three dimensions of judgment in a different random order on the TMC score sheets, as to not place an emphasis on one particular dimension (Amabile, 1982). The judges were informed that the recorded responses were student improvisations, and instructed to rate responses relative to one another rather than against some ideal standard. The two judges, both elementary general music teachers, independently scored the student’s responses and independently discussed details of the scoring procedure with the author throughout two training sessions. The adjudicators rated responses as low (0-2.0 score range), middle (2.1-3.5), and high (3.6-5.0) for each dimension; and were asked to consider their assessments from the pilot study examples as anchors for examining the responses of participants in the main study. The TMC was then administered to 60 participants in the main study, and the two trained judges independently scored the responses.

**Results**

The student improvisation responses to four test items were examined based on originality, rhythmic interest, and melodic interest dimensions. All dimensions were quantified by the independent judges’ ratings. A display of cell means across subscores (originality, rhythmic interest, and melodic interest) and composite scores for the criterion measure are presented in Table 1.

The data also were analyzed through Pearson Product Moment correlations in order to examine the relationship of scores between the two independent judges. Interscorer reliability coefficients were calculated for subscale scores (originality, rhythmic interest, melodic interest) and composite scores on the TMC. The resulting reliability estimates calculated for the TMC subscale and composite scores were: .91 for originality, .87 for rhythmic interest, .83 for melodic interest, and .91 for composite scores. These are all significant at \( p < .05 \).
Table 1. Mean Scores for the Vaughan Test of Musical Creativity

<table>
<thead>
<tr>
<th>Participant</th>
<th>Originality</th>
<th>RI</th>
<th>MI</th>
<th>Composite</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gr. 2 Females (n=10)</td>
<td>10.93</td>
<td>7.72</td>
<td>5.90</td>
<td>24.55</td>
<td>7.27</td>
</tr>
<tr>
<td>Gr. 2 Males (n=10)</td>
<td>10.55</td>
<td>7.10</td>
<td>5.23</td>
<td>22.88</td>
<td>4.52</td>
</tr>
<tr>
<td>Grade 2 Total</td>
<td>10.74</td>
<td>7.41</td>
<td>5.56</td>
<td>23.71</td>
<td>5.96</td>
</tr>
<tr>
<td>Gr. 4 Females (n=10)</td>
<td>12.78</td>
<td>8.28</td>
<td>6.94</td>
<td>28.00</td>
<td>6.89</td>
</tr>
<tr>
<td>Gr. 4 Males (n=10)</td>
<td>13.65</td>
<td>9.49</td>
<td>6.84</td>
<td>29.98</td>
<td>4.51</td>
</tr>
<tr>
<td>Grade 4 Total</td>
<td>13.21</td>
<td>8.89</td>
<td>6.89</td>
<td>28.99</td>
<td>5.76</td>
</tr>
<tr>
<td>Gr. 6 Females (n=10)</td>
<td>11.55</td>
<td>8.35</td>
<td>6.38</td>
<td>26.28</td>
<td>7.78</td>
</tr>
<tr>
<td>Gr. 6 Males (n=10)</td>
<td>11.43</td>
<td>8.13</td>
<td>6.14</td>
<td>25.70</td>
<td>6.10</td>
</tr>
<tr>
<td>Grade 6 Total</td>
<td>11.49</td>
<td>8.24</td>
<td>6.26</td>
<td>25.99</td>
<td>6.81</td>
</tr>
<tr>
<td>Total Females</td>
<td>11.75</td>
<td>8.12</td>
<td>6.41</td>
<td>26.28</td>
<td>7.21</td>
</tr>
<tr>
<td>Total Males</td>
<td>11.87</td>
<td>8.24</td>
<td>6.07</td>
<td>26.18</td>
<td>5.75</td>
</tr>
</tbody>
</table>

Note. RI = Rhythmic Interest subscale, MI = Melodic Interest subscale, and SD = Standard Deviation for Composite scores

To determine whether gender or grade level affected the composite scores, data were analyzed using a two-way analysis of variance with gender (female, male) and grade level (2, 4, 6) serving as the between-subjects factors. The results of that analysis appear in Table 2. Mean composite scores for music improvisational skill ranged from 22.88 (grade two males) to 29.98 (grade four males). There was no significant difference between the mean score for male students (26.18) and the mean score for females (26.28). Although there was no significant gender by grade level interaction; a significant grade level main effect emerged for music improvisational skill scores, $F(2, 54) = 3.52$, $p = .04$. Post-hoc comparisons of grade level scores using a Tukey test revealed that the composite mean score for grade 2 students (23.71) was significantly lower than that of students in grade 4 (28.99), and lower than those in grade 6 (25.99). The difference in mean scores between grade 4 students and grade 6 students was not significant.

Table 2. Analysis of Variance for the Vaughan Test of Musical Creativity Composite Scores

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>$p$</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>0.13</td>
<td>0.00</td>
<td>.96</td>
<td>.00</td>
</tr>
<tr>
<td>Grade</td>
<td>2</td>
<td>140.00</td>
<td>3.52*</td>
<td>.04</td>
<td>.12</td>
</tr>
<tr>
<td>Gender x Grade2</td>
<td>17.53</td>
<td>0.44</td>
<td>.65</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>S within-group error</td>
<td>54</td>
<td>(39.84)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Values in parentheses represent mean square errors. S = subjects.

*p < .05.
Analysis of variance results also revealed a significant grade level effect for originality subscale scores $F(2, 54) = 4.50$, $p = .02$ and melodic interest subscale scores $F(2, 54) = 3.72$, $p = .03$; while rhythmic interest subscales scores were nonsignificant $F(2, 54) = 2.97$, $p = .06$.

The probability of detecting a significant interaction (power) was calculated based on the post-hoc procedures by Cohen (1988). Given a power value of $.30$, the Type II error (1-power) was estimated at $.70$. The present researcher holds a fair expectation of detecting a grade level effect assuming an alpha level of $.05$ and $N = 60$ in this exploratory study. Future studies should utilize larger sample sizes to increase statistical power and precision of results. An effort was made in this study to limit the possibility of Type II error by minimizing some of the extraneous variability through careful random selection of participants, uniform testing conditions, and standardized set of test administration instructions (see Linton and Gallo, 1975).

**Discussion**

The main purpose of this study was to compare the music improvisational creativity of students in grades 2, 4, and 6. The author also investigated whether music improvisational skills could be reliably measured, and if there would be significant gender differences for test scores.

The results of this investigation suggest that music improvisational skill may vary with the grade level of students. A significant grade level difference emerged for test scores with grade 2 students scoring significantly lower than grade 4, and slightly lower than grade 6 students. The results indicate a growth stage may exist from grade 2 to 4, followed by a developmental leveling in selected improvisation skills between grade 4 and grade 6. The significant increase in scores between grades 2 and 4 in the present study supports the results of a series of developmental studies by Brophy (1999, 2002a, 2005) and Kiehn (2000). Brophy proposes a dynamic stage of improvisational development between ages 6 and 9, and Kiehn also points to a growth stage in improvisational skills from grade 2 to 4 (ages 7 to 9) respectively. The results of the present study coincide with both Brophy’s (1999) and Kiehn’s (2000) descriptions of a developmental leveling in improvisation abilities among grade 4 to 6 students.

The present author cautions that little generalizing is possible from such a small sample size. Limited growth in improvisational skills from Grade 4 to 6 may be due to sociological or psychological factors. Kiehn (2000, 2003) discusses that intermediate grade students may feel pressure to conform socially, and if being creative is unique, then individuals may shy away from engaging in creative behavior. Perhaps the appearance of a developmental leveling in improvisational skills is due in part to curricular aspects. Teachers working with intermediate-grade students, for example, may be primarily concerned with covering information, concepts, or rehearsing performing skills that often draw on convergent thinking abilities. Music educators in the present study have approximately 60 minutes of class time per week to teach a wide range of music skills, and one may wonder if this is adequate time for teaching a comprehensive curriculum.

Overall, the grade 2 students in this study improvised with repeated rhythmic and melodic patterns, less adherence to pulse, and contained little evidence of phrasing or motive development. Students experimented with different instrument sounds and combinations of sounds, characterizing the earliest or first level of Kratus’s (1991; 1996) improvisational skill developmental model, exploration. This theoretical model has seven levels and exploration is described as pre-improvisatory behavior. Kratus explains that levels unfold sequentially and build upon each other. A closer examination of the participants’ responses reveal that the grade 4 and 6 improvisations generally adhered to a steady beat, contained developed rhythmic motives, rhythmic antecedent/consequent phrases, and melodic antecedent/consequent phrases. To the listener, melodic motives rarely developed among the improvisations, though surprises seemed to occur when students explored ways to contrast or vary their musical ideas. Familiar tune
fragments (such as “Hot Cross Buns”) were sometimes improvised with the bourdon, which may be viewed as attempts to explore melodic or structural possibilities. The children’s gravitation toward musical convention and creating contrast bear some comparison to the second and third levels of Kratus’s model, process and product orientation. Once the student has explored possibilities, she begins to produce more cohesive patterns (process-oriented) and becomes conscious of structural principles such as tonality and rhythm (product-oriented). Although these responses characterize levels of the Kratus model and findings by Brophy (1999), theoretical models need to be tested further in various settings where students are provided extended time for creative responses.

In this study, no significant relationship was found between gender and improvisation skill. In general, male students have scored significantly higher than females (Kiehn, 2000; Schmidt and Sinor, 1986; Webster, 1977), or no significant differences have been found for gender (Auh, 1995; Baltzer, 1988; Brophy, 1999, 2002a; Swanner, 1985). Several authors found boys scored significantly higher on originality and fluency dimensions of creative thinking (Kiehn, 2000; Schmidt and Sinor, 1986; Webster, 1977). Kiehn suggests gender differences may or may not emerge when researchers consider other dimensions such as elaboration, flexibility, or musical elements. There was no significant difference in originality or in dimensions based on musical elements (melodic interest and rhythmic interest) between females and males in the present investigation. Furthermore; extended response time in the researcher-modified TMC may have enabled participants to be more inventive with their improvisations.

**Implications for Music Education**

Wallach and Kogan (1965) note that extended responses with creative thinking tasks are key, stating “in order really to discover how able a person is at the production of unique associates, it is necessary to allow the person as much time as he wants” (p.18). Music educators may consider expanding lessons to allow students extended time for responses during classroom improvisation activities. Teachers should encourage their students to improvise in many ways, different genres, and on multiple instruments.

Music educators who wish to integrate improvisation activities in their classrooms must address problems with limited instructional time, class size, class schedules, and other issues. Orman (2002) examined 30 general music teachers use of instructional time and reports that improvisation activities comprised only 3% of the classroom curriculum. Arguably, improvisation and composition activities probably require more time to adequately plan and develop. However, the literature indicates that when a teacher provides students with the readiness and skills to create their own music, music becomes the property of the students themselves and meaningful (Daignault, 1996; Guilbault, 2004). Perhaps a direction may be to provide preservice and inservice teachers with more research-based training in improvisation pedagogy and planning.

**Recommendations for Further Research**

Although Vaughan (1971, 1973) reported reasonable reliability coefficients for the TMC (.67 to .90), she discussed modifying the scoring criteria for improved clarity in future research. Additionally, Kiehn (2000) states the judges reported difficulty determining phrase lengths in relation to the fluency subscore for the TMC. Subsequently, the author of the present study aligned scoring criteria to studies in the recent literature (Brophy, 2005; Priest, 2001; Webster, 1990). Acceptable interscorer reliability coefficients ranging from .83 to .91 were obtained in the present study, and this is encouraging given the subjective nature of scoring creative thinking measures.
It appears reliability may have been enhanced in the present study through modification of the scoring procedures, adjudicator training sessions, or having elementary general music teachers independently score student responses to the TMC. The literature suggests that additional practice scoring responses during the pilot phase may be a scoring consistency factor (Baltzer, 1990; Kiehn, 2003). Furthermore; Hickey (2001) and Priest (2006) report that the most reliable group of experts to judge the creativity of students’ musical compositions was elementary general music teachers. Elementary general music teachers were utilized as independent judges during both the pilot phase and main study phase of this project. However, with regard to the limitations of the present study, several validity aspects should be considered in future research.

Content validity could be established for the TMC in future studies through an expert panel and modifications of the instrument. Interestingly, is whether the Vaughan test measures much beyond tendencies in children’s potential to develop music improvisational skills. The open-ended task format of the modified Vaughan test is similar both in approach to other creative thinking tests, and aligns with generally accepted definitions of divergent thinking ability found in the literature (Guilford, 1957; Torrance, 1974, 1998; Webster, 1990). Although the test content has been modified in recent studies, validity could be enhanced by a panel of music educators and theorists meeting periodically to review each item and make further recommendations. An additional open-ended activity asking participants to create a sound story with their voice or other instrument, and a test item requiring students to improvise melodies over a given chord progression or root bass movement could further the align the test content with the National Standards, establishing the test’s usefulness for further research in a variety of grade level settings. Guilbault (2004) discusses the advantage for educators to investigate the type and sequence of root-melody accompaniments most helpful in cultivating children’s improvisational ability; so perhaps future researchers may consider student improvisational responses to the tonic and dominant, or tonic and subdominant chord progressions.

The music education literature on the measurement of improvisation skill reveals that studies involving elementary-age subjects have relied on a variety of instruments and typically utilized small sample sizes with less than 50 subjects. Future investigations with larger samples, reliable instruments, and in diverse school settings appear necessary to generate greater variance in scores and to further substantiate the findings of this study.

There appears to be an immediate need for more investigations utilizing longitudinal and cross-sectional designs, since few studies exist at the elementary and secondary school level. Exploratory studies beyond the sixth grade (age 12) appear necessary to provide a view of how creative thinking develops into adolescent years. Research problems could further address relationships between creative thinking and attitudes, personality traits, or conditions. Although findings in the literature have helped define the process and ways of cultivating creative thinking in music (see Burnard, 2000; Flohr, 1985; Hickey and Webster, 2001; Kratus, 1991; Madura, 1999), we possess little information about current best instructional practices from the teaching field. Additional systematic exploration of the creative thinking process appears necessary to further inform music educators’ instructional theories and practice.
References


