How Different Is Good? How Good Is Different?: The Assessment of Children's Creative Musical Thinking

Maud Hickey

Music Education

Northwestern University School of Music

711 Elgin Road, Evanston, Illinois 60208 USA

mhickey@northwestern.edu

Scott D. Lipscomb

Music Education & Music Technology

Northwestern University School of Music

711 Elgin Road, Evanston, Illinois 60208 USA

lipscomb@northwestern.edu

1 Introduction

You've just collected your fifth grade students' MIDI[1] compositions and, with a hot cup of coffee in hand, are settled down and ready to listen to them from your computer. The assignment for the children was to compose a song on the synthesizer using notation software on the computer. The song was to be eight measures long, in 3/4 time and in the key of Bb. You emphasized that students use Bb as their "home tone;" that is they were to use that pitch at least

3 times in the composition and also to end on the Bb. Your purpose, as a teacher, was to teach about and reinforce the concept of key centeredness (i.e. tonality), as well as to determine whether the students understood 3/4 time. A second—though no less important—purpose for this assignment was to give children the chance to be creative in their approach to learning. You smile and nod as you listen to the first ten or so compositions, all just a little different, but mostly the same: clearly following the parameters that you set to create a simple, single line melody. But when you get to Nora's song you are startled. Though she *did* write in 3/4 time and used the Bb as asked, she clearly experimented with several different timbres and composed a jagged atonal melody full of wide leaps, accompanied by alternating loud/soft tone clusters using an electronic-sounding timbre. It didn't sound very "good" to you, but yet it was somehow interesting. Was it a random mess? Or did Nora compose this song purposefully? How should it be graded? How do you respond to Nora? It certainly was not nearly as "neat" and tonally "centered" as the other student compositions. In fact, it was downright strange. You're stuck with these questions, yet also intrigued by what Nora composed.

Every music teacher who has incorporated any kind of composition in the music classroom has undoubtedly experienced something similar to the imaginary scenario described above. For those who typically give less structured, more "open" assignments (with virtually no parameters), the percentage of "peculiar" sounding compositions is even greater. Upon first hearing, the most unusual compositions may be dismissed as "wrong," or "not following the rules," or simply "bad." Our music teaching culture tends to favor the "safe" side—that is providing structure in composition tasks in order to assure that students create something that sounds "good." Teachers feel more confident assessing the more structured, neat, "tonal," approaches to music creation;

especially if they have not been trained formally in music composition. Yet experimentation and novelty are the sine qua non of creativity. How can we facilitate student learning and creative use of both the worlds of rule-bound composing and free creativity? What means can we utilize to determine when a child has acquired the ability to combine these worlds?

What constitutes a *good* composition? What constitutes a *creative* composition? Where does "highly unusual" fit in? Can different be good? How can good be different? To answer these questions and address the issues posed above we will examine approaches to assessment in creativity and in ethnomusicology and share studies in which the present authors have applied these approaches to the assessment of elementary children's musical compositions.

2 Approaches to the assessment of creativity

As the "grandfather" of creativity assessment, J. P. Guilford's long quest to measure creativity began with his 1950 address to the American Psychological Association (Guilford 1950). Guilford's Structure of Intellect (SOI) model proposes 180 cells of thinking operations. Thirty of these cells fall under divergent production abilities which Guilford purported as important to creativity (1967, 1988). Tests that measure creativity based on the SOI model measure the variables of fluency, flexibility, originality, and elaboration. The Torrance Tests of Creative Thinking (Torrance 1974) are the most widely used standardized tests of creative thinking that emerged from Guilford's SOI model.

In music, Webster (1994) adapted these four factors to create the *Measurement of Creative Thinking in Music* (MCTM). It is probably the most well known and thoroughly researched tool for assessing creative thinking in music. In the MCTM, the student is prompted to perform a series of improvisations based on imaginative scenes, such as a robot in a shower, a frog jumping on lily pads, or a rocket launching into space. The student responds to these prompts using a foam ball on a keyboard, their voice in a microphone, or temple blocks. The resulting musical improvisations are recorded and scored for extensiveness, flexibility, originality, and syntax, as well as for an overall musical creativity score.

The foci in both the Torrance and Webster approaches are to rate the overall creativity, or creative thinking ability of the test taker based on the premises that creativity can be measured through test exercises, and is based on the factors of fluency, flexibility, originality and elaboration. For the purposes of this paper, we are interested in observing the creativity of children's music compositions, and examining the efficacy of social methods for measuring these.

2.1 Creative product

A widely held definition of a creative product is that it is both "novel" and "appropriate" (Amabile 1983, Baer 1997, Davis 1992, Mayer 1999). Of course "novel" and "appropriate" can and do have a variety of meanings depending upon the context. A main consequence of this definition is that a product that is *only* original without any sense of appropriateness or

usefulness in the culture is not creative, and vice versa; a product that is appropriate or valuable without any degree of originality is not creative.

What we find to be a very useful definition for creative products when dealing with children is that offered by Baer (1997): "Creativity refers to anything someone does in a way that is original to the creator and that is appropriate to the purpose or goal of the creator" (p. 4). This definition supports what some call "small c" creativity (Feldman *et al*. 1994, Gardner 1993) whereby every person is more or less "creative," and the more or less is in comparison to others in their cultural and social context. For children in a classroom, then, the most creative products are those that are the most unusual, yet appropriate, in the context of that classroom or age-group within that cultural milieu. "Appropriate," in this context, means aesthetically interesting (this might be pleasing or not pleasing; simply catchy or interesting). A musical composition for a 10-year-old child that is considered "creative" will be both interesting as well as novel or unusual in comparison to others in her age group. Nora's composition described in the opening scenario would fit into this category.

2.2 Consensual assessment

Amabile (1983) devised a "consensual assessment technique" (CAT) for rating the creative quality of art products and which aligns with the definition of creativity described previously. The technique is based on her consensual theory of creativity, suggesting that creative ability is best measured by assessing the creative quality of the products that are a result of creative endeavors. Furthermore, Amabile proposed that subjective assessment of such products by

experts in the domain for which the product was created is the most valid way to measure creativity. Amabile argued that it is not possible to articulate objective criteria for a creative product. Rather, she asserts:

> A product or response is creative to the extent that appropriate observers independently agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated. Thus, creativity can be regarded as the quality of products or responses judged to be creative by appropriate observers , and it can also be regarded as the process by which something so judged is produced.
>
> (Amabile, 1983, p.31)

Amabile (1983) lists necessary conditions and requirements regarding the creative tasks and methods for successful utilization of the consensual assessment technique. Three requirements must be met in selecting an appropriate task: (1) the task must result in a clearly observable product or response that can be made available to appropriate judges for assessment, (2) the task must be open-ended enough to permit flexibility and novelty in response, and (3) the task should not depend heavily on special skills that some individuals may have developed more fully than others.

Amabile (1996) reports – by author, task/product, subjects, and judges used – the results of approximately 53 different studies that utilized the consensual assessment technique for rating creativity in a variety of artistic domains (visual art, poetry, and story telling). Inter-rater reliability scores for the reported studies are consistently high. Several researchers have utilized or tested the CAT in visual art, and in poetry and story writing, also with consistently high inter-rater reliability, supporting the construct validity of this technique.

The CAT has been modified and used successfully for rating the creativity of musical compositions by Bangs (1992), Hickey (1995), Daignault (1997) and Brinkman (1999), and for rating musical improvisations by Amchin (1996) and Priest (1997, 2001).

While the consensual assessment technique assumes that "expert" judges can reliably rate creative products, recent research has examined who the best "experts" might be. Runco *et al*. (1994) sought to determine which group of judges was most reliable for judging the creativity of visual artwork when using consensual assessment. College level subjects created 3 artworks to be self-rated, rated by peers, and rated by professional artists for creativity. The self-assessment rankings and peer assessments rankings for subjects' art works were similar, while professional judges also correctly ranked the drawings, but the differences between rankings were not significant and the scores given by the professionals were much lower than those given by the students.

Hickey (2000) sought to find the best group of judges when using a consensual assessment technique to rate the creativity of children's music compositions. She compared the reliability of creativity ratings of 10-year-old children's original musical compositions among different groups of judges. The inter-rater reliabilities for each group's creativity ratings were: .04 for composers; .64 for all music teachers combined, .65 for instrumental music teachers; .81 for general/choral teachers; .70 for music theorists; .61 for seventh-grade children; and .50 for second-grade children. Hickey suggested that maybe the best "experts" for judging creativity are not those who

are professionals in the field, but those closest to the students who are creating the works (in this case, teachers).

Webster and Hickey (1995) compared the reliability of open-ended ("consensual assessment" type) scales to more closed, criterion-defined scales for rating children's musical compositions and/or creativity. They discovered that rating scales using consensual assessment as outlined by Amabile were at least as reliable – if not more reliable than – scales with more specific criterion items (see Figure 1).

The CAT provides a method for researchers to identify creative musical compositions of children in a realistic and valid manner. It conforms to the widely held social definition of creativity and supports "small-c" creativity. While teachers are not likely to use this method as a form of assessment in their classroom, the premise upon which it is based can help teachers understand that "unusual" can be good. In fact, "unusual" might even signify creative potential in a given child. Music research incorporating the CAT also confirms that music teachers do have the ability to correctly identify varying levels of creativity as evidenced in the compositions of children.

Figure 1: Rating Scale Samples from Webster & Hickey (1995).

Specific Musical Characteristics (presence)

Rhythm

• The degree to which the composition shows a pleasing use of rhythm.

5 4 3 2 1

Texture

• The degree to which the composition shows a pleasing use of texture
(use of more than one instrument or pitch at a time)

5 4 3 2 1

Global Considerations

• First impression

5 4 3 2 1

• Imaginative varying and ornamenting

5 4 3 2 1

• In general, the degree to which the composition has aesthetic value

5 4 3 2 1

3 Cantometrics

3.1 Background

Because music is a cultural artifact and, as a result, musical creativity must be considered within a cultural context, we turn our attention to a method of analysis developed specifically for that purpose. In the study of music "as a form of human behavior," Alan Lomax (1962, p. 425; see

also Lomax 1976 and Nettl, 1964) has been one of the most prolific researchers in the field of ethnomusicology. He developed the system of "cantometrics" which, using a series of 37 qualitative judgments, "enables a listener to listen to a recorded song from anywhere in the world in a matter of minutes" (Lomax 1962, pp. 428-429). The 37 scales in Lomax' original list can be grouped into meaningful subcategories, including group organization, level of cohesiveness, rhythmic features, melodic features, dynamic features, ornamentation, and vocal qualities (Lomax, 1976 p. 18). Though compositions by student composers undoubtedly emerge from within a social milieu, some of the more creative challenge the rule system, limitations, and constraints imposed by that context. As a result, the application of cantometric analysis to these compositions allows a method of assessment that is not burdened by the assumptions of any single cultural style and does not inherently impose the quality of "good" or "bad" upon a given work. Instead, purely musical traits of the composition – "gross traits rather than the detail of music," according to Lomax (p. 426) – are observed objectively and these ratings are used to compare across compositions. Lomax and an assistant reviewed approximately 400 recordings from 250 different culture areas as a means of testing the viability of cantometrics as a system of analysis (Lomax 1962). Within the context of the present study, the comparisons were, of course, made across compositions, rather than social groups, yet the application of this technique proved highly successful.

3.2 The present study

In the experiment that we will be reporting, a subset of 13 scales were used rather than Lomax' complete set of 37. This decision was made due to the fact that many of the scales would not

have discriminated the compositions to be evaluated, due to the nature of the assignment. The 13 chosen scales, along with the various categorical values for each, are provided in Table 1.[2] For more details about the scales and their application in this analytical context, consult Lipscomb *et al*. (in press).

Student compositions analyzed for this study were taken during the fourth week of a 10-week Creative Music Project. Fifth grade (ages 9- and 10-years-old) students from four music classes (N=86) at Monroe May Elementary School in San Antonio, Texas participated in this study. A grant from Texaco Corporation afforded the opportunity to purchase SoundBlaster Live! sound cards, LabTec LT 835 stereo headphones, and BlasterKey keyboards for each of the 25 computer stations in the lab. The 10-week project consisted of a tonality judgment pre-test, eight weeks of instruction in compositional techniques, and a tonality judgment post-test. Taught by Dr. David. Sebald (UTSA), the instructional component of the study focused primarily on musical form, but also introduced other musical elements as a means of introducing the concept of musical organization (e.g., rhythm, meter, tempo, texture, harmony, melodic repetition, contour, etc.). Students were also instructed in the basic use of Cakewalk Express, a MIDI sequencing program, as a means of recording their musical ideas. The present chapter will focus on the cantometric analysis of student compositions collected midway through this instructional process.[3]

Table 1: The 13 cantometric scales used in the present study; selected and modified from the list of 37 used by Lomax (1962). A category of "NA" (not applicable) was added in some cases.

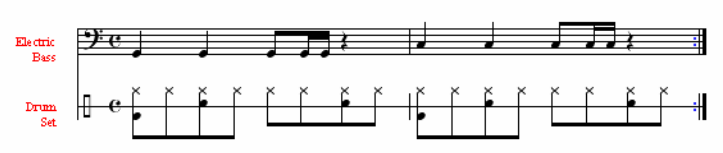
Qualitative scales used:
1. musical organization of instruments (musical texture) no instrument – monophonic – unison – heterophonic – homophonic – polyphonic

2. rhythmic coordination of instruments (blend)
   little to none – minimal – good – unison – maximal
3. overall rhythmic structure (meter)
   free – irregular – one beat – simple – complex
4. melodic Shape (contour)
   NA – arched – terraced – undulating – descending
5. musical form
   through-composed – repetitive with variation – repetitive without variation – strophic – canonic – other
6. phrase length (number of measures)
   more than 8 – 5 to 8 – 3 to 4 – 2 – 1
7. number of phrases
   more than 8 – 5 to 7 – 4 or 8 (symmetrical) – 4 or 8 (asymmetrical) – 3 or 6 (symmetrical) – 3 or 6 (asymmetrical) – 2 (asymmetrical) – 1 or 2 (symmetrical)
8. position of final tone
   NA – lowest tone – lower half – midpoint – upper half – highest tone
9. keyboard range
   within P5 – within octave – 1 to 2 octaves – 2 to 3 octaves – >3 octaves
10. dominant melodic interval size
    NA – monotone – <=semitone – whole step – maj/min 3rd – P4 or larger
11. polyphonic type
    none – drone – isolated chords – parallel chords – harmony – counterpoint
12. use of tremolo
    little or none – some – much
13. use of accent
    unaccented – some – main pulses – main beat pattern – most notes

Two specific research questions guided this research. First, can typical students learn to create music effectively with the technologies (i.e., computer, sequencing software, MIDI keyboard, etc.) described above? Second, can Lomax' "cantometrics" (1962, 1976) provide a reliable method for analyzing these student compositions? Each investigator independently evaluated the 86 student compositions in two ways: using 13 cantometric scales and on a scale of dissimilarity in reference to a "standard." For the specific composition assignment being evaluated, students were given a repeating two-measure percussion beat pattern (Figure 2) and were free to incorporate, edit, vary, and/or use this building block in any way they saw fit in the process of creating their composition. For the dissimilarity judgments, the original repeating two-measure

rhythmic pattern was used as the standard, affording an opportunity to judge how much a given student composition varied from the material initially provided to each student by the instructor. Inter-rater reliability was very high for both the cantometric scales (r = .82) and the dissimilarity ratings (r = .80).

Figure 2: The two-measure rhythmic sequence provided to students as a basis for their musical composition.



In the following presentation of cantometric ratings, we will discuss two groups of students: those whose compositions were judged to be "most different" in the dissimilarity rating task and those whose compositions were defined as "more similar" (i.e., less dissimilar). The former group was operationally defined as any individual whose composition received an average rating of 4.5 or greater on the scale of dissimilarity ("1" = most similar; "5" = most dissimilar) in comparison to the standard. Obtaining such an average required that either one or both of the investigators assign a rating of "5." Of the seven compositions included in this category, five were assigned a rating of "most dissimilar" by both investigators, while the remaining two compositions received a rating of "5" from one investigator and "4" from the other. When a cantometric profile was created to compare these two groups – "different" (n = 7) and "more similar" (n = 79) – notable differences emerged. A visual representation of these profiles is provided in Figure 3 and a brief verbal description of the most notable differences is provided in Table 2. In accordance with Lomax' instructions, the profiles in Figure 3 were created by identifying the category within each scale that represented the most frequent occurrence within

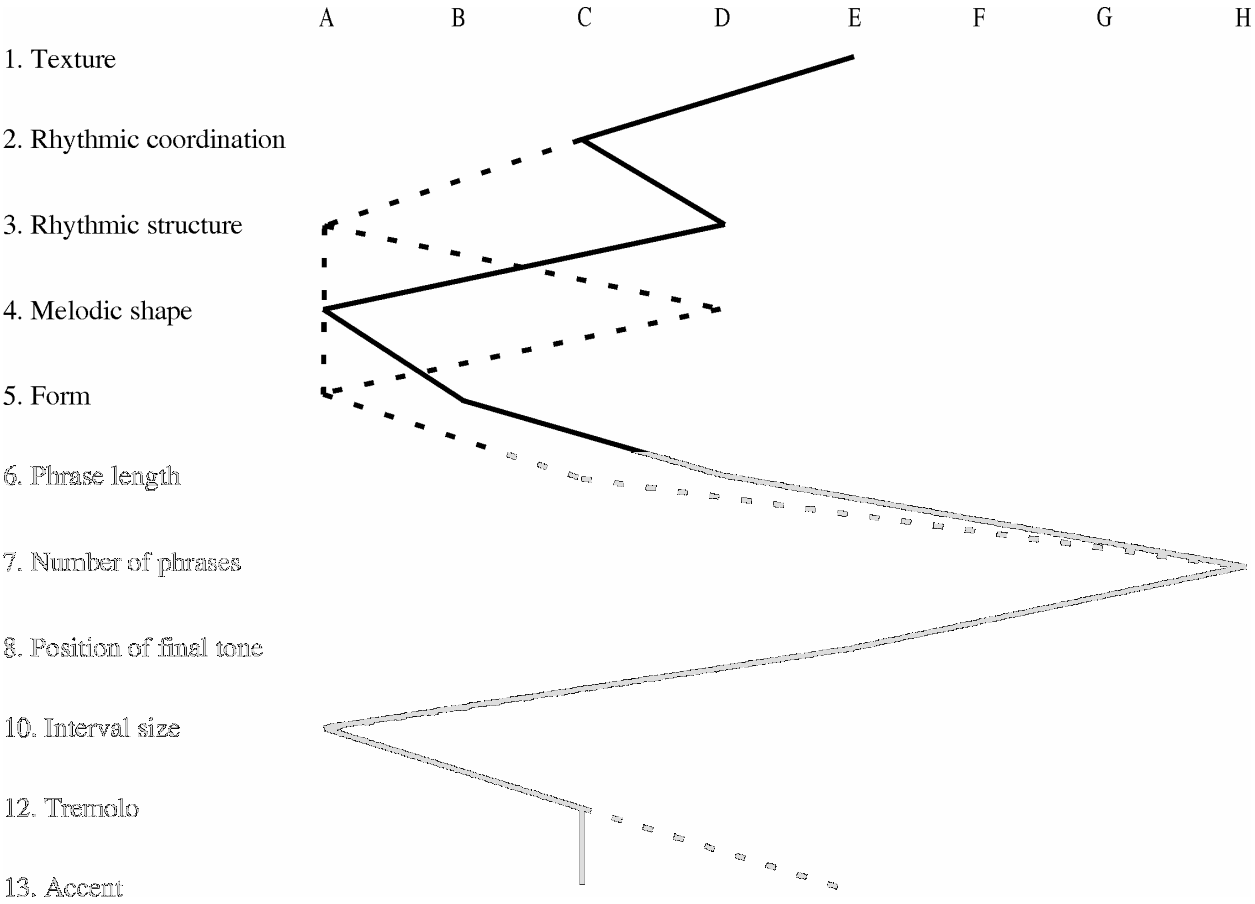
the group. These "most frequently occurring categories" are then connected by a line from one scale to the next. In the figure, a dashed line represents the profile for the "different" group (D), while a solid line represents the profile for the "more similar" group (MS).

As one can instantly perceive from the differential profiles in Figure 3, students whose compositions were rated "different" in comparison to the standard appear to have utilized different compositional strategies than the other "more similar" group. The most substantial differences are identified in Table 2. It is, perhaps, no surprise to find that the greatest number of differences occur in the manner in which melodic features are manipulated. Rhythmic and dynamic features (e.g., accent)[4] also play an important role in this distinction.

Table 2: Comparison of selected item differences between all compositions and compositions from "most different" group.

Item:	"Less Different" Compositions	"Most Different" Compositions
3. rhythmic structure	Choice D (simple) 95%	Choice A (free) 43%; Choice D (simple) 36%
4. melodic shape	Choice A (no discernable melody) 90%	Both Choice A (no discernable melody) and Choice D (undulating) 38%
5. form	Choice B (repetitive with some variation) 43%	Choice A (through-composed) 50%
6. phrase length	Choice D (short 2 ms.) 64%	Choice C (medium 3-4 ms.) 54%
10. interval size	Choice A (no discernable melody) 90%	Choice A (no discernable melody) 38%; Choice C (1/2 step or less) 31%; Choice F (4ths & 5ths or larger) 15%
13. accent	Choice C (medium, accents conform to main beat pattern) 57%	Choice A (very forceful) 21%; Choice D (relaxed) 21%; Choice E (very relaxed) 43%

Figure 3: Overlaid Cantometric profiles for "more different" (dashed line) and "more similar" compositions (solid line). The letters (A to H) at the top of the figure refer to the various responses to each given scale provided in Table 1. The first potential response is represented by "A," the second by "B," etc. [Note that the triangle shape around item 4 (melodic shape) results from the fact that an equal number of the compositions fell into categories A and D.]



Almost all MS compositions (95%) were identified as "simple" when evaluating their rhythmic structure. In contrast, the D compositions revealed a higher degree of complexity and variety. Though many of these compositions were also categorized as "simple" (36%), many were assigned to the "free" category (43%). Concerning the presence of accent, the MS compositions were categorized primarily as "medium," described as conforming to the main beat pattern. Interestingly, very few of the D compositions were assigned to this middle-ground category. Instead, there was significant variability in the way that accents were either present or not: very

forceful (accents falling on most notes; 21%), relaxed (some accent; 21%), and very relaxed (nearly unaccented; 43%). It appears that, though a small percentage of students in the D group used forceful accents, this rhythmic aspect of musical composition subdued in comparison to the MS group.

A large proportion of the MS group (90%) utilized no discernable melody in their composition. This may not be as surprising as it seems at first, given that the template provided to each student contained only a basic drum rhythm and bass line. The addition of a melodic component required a creative leap on the part of the student composer. A substantial group of the D group compositions (38%) were also evaluated in this same category. However, an equal number of compositions (38%) were categorized as "undulating," meaning that not only did these students add a melody to their composition, but they also created a coherent up-and-down melodic contour. The dominant melodic interval also revealed a significant difference between the groups. Though the same percentages were categorized as "no discernable melody" (90% for MS and 38% for D), the D group revealed a greater range of variability. In fact, 31% of the compositions used a dominant interval size of a half step or less, resulting in a highly chromatic melodic context. Another small but significant proportion (15%) utilized mostly perfect fourths and fifths.

When considering overall musical form, compositions in the MS group tended to fall into the "repetitive with some variation" category (43%), an organizational structure familiar to all students from the many familiar folk melodies and daily listening to popular music forms. In dramatic contrast, 50% of the D group submitted compositions that were categorized as

"through-composed." Phrase lengths also differed between the two groups. Compositions by the MS group consisted of short two measure phrases (64%), while the majority of D compositions exhibited phrases that were three to four measures in length (54%).

In conclusion, the use of cantometrics as an evaluative tool allows us to determine that those compositions considered "most different" from the standard template provided by the instructor evidence certain musical traits that distinguish them from those compositions that are "more similar" to the standard. Specific musical characteristics that differentiate these groups of compositions include:

- freer rhythmic structure
- examples of heavily accented and nearly unaccented compositions, rather than the middle ground use of accent evidenced in compositions of the MS group
- the innovative addition of an undulating melodic contour to the rhythmic underpinning provided by the musical template
- the dominant use of small (semitone) and large (perfect fourths and fifths) melodic intervals
- through-composed musical forms, rather than thematic variation
- longer phrase lengths

4 Further research

The study reported above opens the door to a wide range of research possibilities. Lomax' cantometric system has proven quite useful in determining perception-based differentiation

between student compositions. More research is needed to determine its viability and additional contexts within which this method may prove of use.

Further research is needed to continue to examine the validity of the consensual assessment technique, and to compare it to Webster's MCTM. In addition there is a need to examine the connection between the process of children's creative musical thinking and the creative success of their final compositions in order to help teachers support this success in their classrooms. How might either the consensual assessment technique or Webster's MCTM be used to view this connection between process and product?

Finally, it is worth noting that the goal of this research was *not* to evaluate student compositions in regard to some standard of "quality." Instead, we wanted to identify specific differences between student compositions for use as a means of considering the various ways in which students approach such a creative task. The assessment of quality – whatever that might mean in the context of student compositions – remains, as yet, unmeasured.

5 Conclusions

Two questions we posed at the beginning of this chapter asked: How can we facilitate student learning and creative use of both the worlds of rule-bound composing and free creativity? What means can we utilize to determine when a child has acquired the ability to combine these worlds? By identifying and then examining a group of children's compositions using the Cantometric lens created by Lomax, we were able to identify those most "different," and delineate the

characteristics of these composition. We hope by understanding that different *can* be good (and easily identified), that teachers support and even encourage compositions that use free rhythmic structure, through-composed musical forms, innovative melodic use and longer phrase lengths than might be typical or expected for elementary-grade children. Composition assignments should be balanced between structure and freedom in order to facilitate children's growth in free creative thinking. We need to be sensitive to the odd compositions that are created by children and not dismiss them immediately as "wrong," but rather embrace the thinking that pushes the norm.

What constitutes a *good* composition? What constitutes a *creative* composition? Where does "highly unusual" fit in? Can different be good? How can good be different? The present authors believe that different *is* good, and good *is* different when it comes to children's compositions. If as teachers we want to *encourage* creativity, then we should support and promote that which might be perceived as "different." While it is certainly true that rules, theory, and basic musical skills form an important part of music instruction, it is important for teachers to realize that compositions that sound "different" do not necessarily constitute "bad" music. This realization will allow students to produce truly creative work – even that which is conceived as extreme – and will not act to censor students whose creative output is "different" from the norm. It is quite possible that such an individual has provided evidence of unusual creative potential. In order to capture such creative potential, in fact, it may prove useful at times to evaluate as "positive" not how closely the results of a student's creative effort fit within the confines of a guided assignment, but how far beyond the boundaries the student can go while still producing a unique, yet coherent, creation.

6 Acknowledgements

We would like to acknowledge the supporting agencies who made this research possible: Northwestern University, The University of Texas at San Antonio, May Elementary School, and Texaco Corporation.

---

[1] MIDI stands for Musical Instrument Digital Interface and is the standard file format that is created using a digital instrument such as a synthesizer/keyboard and music sequencing or notation software.

[2] For a complete list of Lomax' 37 scales in their original form and examples of completed coding sheets, see Lomax (1962, esp. pp. 429-431).

[3] Results of the tonality experiment have been reported elsewhere (Hodges & Lipscomb, 2004; Lipscomb & Hodges, 2002).

[4] Though Lomax places the "accent" scale in the "Vocal Qualities" category, in the context of the present study, the present authors believe it belongs in the "Dynamic Features" category due to both the basic tenets of the Western musical tradition and the manner in which this scale was rated within this analytical context.

Bibliography

Amabile, T. M. (1983). *The social psychology of creativity.* Springer-Verlag, New York.

Amabile, T. M. (1996). *Creativity in Context. Update to the social psychology of creativity.* Westview Press, Boulder, CO.

Amchin, R. A. (1996). Creative musical response: The effects of teacher-student interaction on the improvisation abilities of fourth- and fifth-grade students. (Doctoral Dissertation, University of Michigan, 1995). *Dissertation Abstracts International*, 56/08, 3044A.

Baer, J. (1997). *Creative teachers, creative students.* Allyn & Bacon, Needham Heights, MA.

Bangs, R. L. (1992). *An application of Amabile's model of creativity to music instruction: A comparison of motivational strategies*. Unpublished Doctoral Dissertation, University of Miami, Coral Gables, Fl.

Brinkman, D. J. (1999). Problem finding, creativity style and the musical compositions of high school students. *Journal of Creative Behavior*, 33(1), 62-68.

Daignault, L. (1997). Children's creative musical thinking within the context of a computer-supported improvisational approach to composition. (Doctoral dissertation, Northwestern University, 1996). *Dissertation Abstracts International*, 57, 4681A.

Davis, G. A. (1992). *Creativity is forever* (3rd ed.). Kendall/Hunt Publishing Company, Dubuque, IA.

Feldman, D., Csikszentmihalyi, M., & Gardner, H. (1994). *Changing the world. A framework for the study of creativity.* Praeger, Westport, CT.

Gardner, H. (1993). *Creating minds*. Basic Books, New York.

Guilford, J. P. (1950). Creativity. *American Psychologist*, 14, 444-454.

Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill, New York.

Guilford, J. P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, 48, 1-4.

Hickey, M. (1995). *Qualitative and quantitative relationships between children's creative musical thinking processes and products.* Unpublished Doctoral Dissertation, Northwestern University, Evanston, IL, 1995.

Hickey, M. (2000). The use of consensual assessment in the evaluation of children's music compositions. *Proceedings from the Sixth International Conference on Music Perception and Cognition* (ed. C. Woods, G. Luck, R. Brochard, F. Seddon, and J. A. Sloboda) [CD-ROM], Keele, UK, August, 2000.

Hodges, D. & Lipscomb, S.D. (2004). Tonality judgments in popular music contexts by preteens and college students: A comparative analysis. Manuscript in preparation.

Lipscomb, S.D., Hickey, M., Sebald, D., & Hodges, D. (in press). The Creative Music Project: A cantometric analysis of fifth grade student composition. *Journal of the Centre for Research in Education and the Arts.*

Lipscomb, S.D. & Hodges, D. (2002). Tonality judgments in popular music contexts by preteens and college students: A comparative analysis. Paper presented at the Music Educators National Conference. Nashville, TN (April, 2002).

Lomax, A. (1962). Song structure and social structure. *Ethnology*, 1, pp. 425-451.

Lomax, A. (1976). *Cantometrics: A method in musical Anthropology*. University of California Extension Media Center, Berkeley, CA.

Mayer, R.E. (1999). Fifty years of creativity research. *Handbook of creativity*, (ed. Sternberg, R. J.), pp. 449-460. Cambridge University Press.

Nettl, B. (1964). *Theory and method in Ethnomusicology*. Collier-Macmillan, Ltd, London.

Priest, T. (1997). Fostering creative and critical thinking in a beginning instrumental music class. (Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign). *Dissertation Abstracts International*, A 58/10, p. 3870, Apr 1998.

Priest, T. (2001). Using creative assessment experience to nurture and predict compositional creativity. *Journal of Research in Music Education*, 49(3), 245-257.

Runco, M. A. , McCarthy, K. A., & Svenson, E. (1994). Judgments of the creativity of artwork from students and professional artists. *The Journal of Psychology*, 128(1), 23-31.

Torrance, E. P. (1974). *The Torrance tests of creative thinking: Technical-norms manual* (revised edition). Scholastic Testing Services, Bensenville, IL.

Webster, P. (1994). *Measure of Creative Thinking in Music-II (MCTM-II). Administrative Guidelines*. Unpublished manuscript, Northwestern University, Evanston, IL.

Webster, P. & Hickey, M. (1995). Rating scales and their use in assessing children's compositions. *The Quarterly Journal of Music Teaching and Learnin*g, VI (4), 28-44.

Webster, P. (1994). *Measure of Creative Thinking in Music-II (MCTM-II). Administrative Guidelines*. Unpublished manuscript, Northwestern University, Evanston, IL.