

Examining Correlations when Using Amabile's Consensual Assessment Technique to Support
Validity of Teachers as Expert Judges

by Patrick K. Cooper

Abstract

The purpose of this paper is to test the reliability of different assessment groups when Amabile's consensual assessment technique is used to score children's musical compositions on a variety of subjective ratings, examine why we get the results we do using the consensual assessment technique, and make the case for a group of teachers as expert judges of creativity using the consensual assessment technique. The research found, through the application of Amabile's consensual assessment technique, that a highly qualified music teacher is a reliable judge of creativity in musical compositions while other groups of peer assessors and random adults were not. This was supported by a correlational study indicating the teachers' ability to assess skill quality that was not found in any other group of assessors.

Introduction

In an environment where educational decision makers are calling for more and more standardized assessments backed by explicitly defined criteria and assessment procedures, music teachers are often hard pressed to find a way to explain their rationale for assessing their students the way they do. Some institutions require music students to be assessed on creativity, for example, to pass one class and reach a higher level (Cowdroy and Williams, 2006; Sheridan and Byrnes, 2002). This can be difficult due to the lack of a standardized definition of creativity in education and how to fairly rate the creativity of a product. One solution is the Consensual Assessment Technique first described by Teresa M. Amabile (1982). She offers a definition of

creativity that “implicitly underlies most subjective creativity assessment methodologies.”

According to Amabile, “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” (p. 998). Amabile’s approach could be a potential solution to assessing creativity in multiple arts domains, particularly at institutions requiring that assessments be made on students’ creative work.

Prior research has shown that music teachers and peer-student groups are reliable raters of creativity using this technique (Amabile, 1982; Hickey, 2001). One study examined the reliability of different assessment groups and found those familiar with the domain in which they were rating scored as reliable raters using Amabile’s consensual assessment technique (Hickey, 2001). However, in research about using this technique there is some information missing that may be useful. I believe it would be beneficial to examine the reliability of the technique among other subjective ratings, and examine the correlations between the ratings. Amabile explains that creativity should not be the only rating being assessed when using the technique (1982). If there are certain correlations between the other ratings, there could be information telling us why certain assessment groups are reliable and why others may agree, but for the wrong reasons.

I will examine why assessment groups rate as reliable and why one group should be considered expert judges and other groups should not despite the fact they all may initially appear to be reliable raters using this technique to assess creativity.

Student Composition Project

The compositions that were assessed in this study were the creations of 6th grade strings students in a middle-class suburb of Phoenix, Arizona. The students were second year musicians that had received approximately 50 minutes of instruction in mixed groups per week in their first year and 100 minutes of instruction in mixed groups per week in the first three quarters of their second year. Students had participated in two compositional projects in their first 1.75 years of instruction prior to this one with the objective of providing sound-effect tracks to poems containing onomatopoeias. This was to be their first attempt at composing without restrictions and the first time composing individual projects.

The instructional time used included a 10-minute introduction and demonstration of the project we were going to be doing, and two 40-minute sessions of time to compose and record products. During the 10-minute introduction, a 2-measure composition in 4/4 involving only quarter notes and quarter rests were drawn on the whiteboard. Time was given for students to list every way they could think of to alter the composition in front of them. Responses included changes in components of meter, modality, variations on rhythms and notes used, techniques such as pizzicato or staccato, dynamics, form elements such as repeats or codas, and abstract techniques such as using the instrument as a percussive instrument or to bow below the bridge. These student-submitted ideas were recorded and included in a handout for the students to reference as they composed. The reverse side of the handout included staff paper for the students to write down their songs in standard music notation; students were encouraged to take new handouts as needed. Students were not required to use standard music notation and were instructed that what was most important is that they have a way to perform their song, whether that meant writing down the song in some form or memorizing it. Students were encouraged to

experiment with sounds and ideas before starting to formally compose. When they were finished, students were offered the chance to record their projects and listen to their final product.

Several precautions were taken to offer students the best possible chance to complete their projects in an environment conducive to creative tasks based on the idea of Csikszentmihalyi's "flow" state. This creative environment was attempted as a means to "initiate and plan innovative and exciting learning opportunities, to monitor, regulate and assess learning" (Sheridan and Byrne, 2002). Sheridan and Byrne suggested attempting to foster the nine components of enjoyment that many people experience when engaged in activities such as sports and hobbies as described by Csikszentmihalyi. Below are the nine components discussed by Sheridan and Byrne (2002) and ways the realization of these components was attempted in this study:

- There is no worry of failure;
 - o The project was not graded and students were not required to complete their project in class or at any particular time. Students were told the recording of compositions were voluntary and for their enjoyment only.
- There are clear goals every step of the way;
 - o The first stated goal was to experiment with sounds and techniques. The second goal was to begin composing. The third goal was evaluate whether they approved of their composition and if they wanted it to be recorded. The final goal was to repeat the process or start a new project with a friend who had also finished.
- There is immediate feedback;
 - o Students often played each other their compositions during the composing process. They chose friends and others they knew would give them feedback in a safe way. Occasionally, a student would come play their composition for me in which case I asked them if they wanted to record or continue revising. They made that judgment and if they were ready they often chose to record.
- Distractions are excluded from consciousness;
 - o At no point during the project did I get up and walk around. Students were observed from afar and no interruptions were made after initial directions were given. In observation, students did not interrupt each other and were eager to work on their projects individually.
- There is a balance between challenge and skill;
 - o This was a new challenge for each student. The amount of challenge presented most likely correlated to their skill level. If a student had a high expectation

for themselves, they most likely had more skill, and most likely set a more difficult challenge, the challenge being writing a composition they were willing to have recorded and share with their teacher.

- Self-consciousness disappears;
 - o Students had received instruction for 1.75 years together and played multiple playing tests individually in front of each other. Students often played each other their compositions and sought immediate feedback. They were comfortable enough doing this, and sharing their projects with me, that I believed the recorded projects were done without the fear of self-consciousness.
- Sense of time becomes distorted;
 - o It appeared students were focused and worked hard in the time they had to complete projects in class. I did not survey the students on this but I would expect a high level to respond that “time flew” while composing.
- Action and awareness are merged;
 - o I interpret this to mean the student becomes aware of how their actions affect the outcome of their product. The student becomes aware that adding pizzicato will accentuate a passage and takes action by adding it to their composition.
- The activity becomes autotelic (worth doing for its own sake)
 - o Because there was no grade assigned to the project, those who completed and asked to be recorded were doing so for autotelic reasons.

Sheridan and Byrne’s (2002) research argued that many of these conditions are significant in the assessment of creative tasks in music. When these conditions are not met, emotions such as anxiety and worry can result, likely stifling creativity. When requiring a certain passing grade, some anxiety and worry may already exist. However, aspects of the project itself such as the lack of restrictions may be essential in fostering an appropriate atmosphere students are expected to produce an accurate representation of their creativity for which to assess.

Setup of the Assessments and Rating Groups

All raters were asked to assess the 12 student-created compositions using a Likert-type 5-point scale. At the suggestion of Amabile’s research, the judges made multiple subjective ratings. The ratings chosen were *I liked it*, *it was creative*, *it was unique*, and *it showed the quality I expect for this skill level*. Each composition was listened to twice and a brief period of time was given between each composition to finish ratings. Groups were asked to score these assessments

using their own definitions of the rating to be made. Attempts by raters to receive clarification or defined criteria were rejected with the exception of the expected skill level quality. All groups were given the information about the students' educational backgrounds and prior performance experiences.

Three groups of raters were chosen to assess the same musical compositions. A group of eight 5th-6th grade teachers evaluated the composition and was made up of string teachers with between three years and thirty-two years of experience. I expected these string teachers to be expert judges of creativity as they all teach the same level in the same district. A second group of 16 6th grade students all came through the first 1.75 years of instruction together by the same teacher. One might expect that this student group to be a reliable rater of creativity, but I would argue for the wrong reasons. Hickey (2001) found that second-grade students have similar scores when asked if they liked the product. Of these 16 students, 4 of the 12 compositions assessed were from students in this very group. The group of 5 non-musician adults consisted of randomly selected adults who agreed to participate in the project without compensation. They acted as the control in this group to see if they have similar correlations in their ratings to the other groups.

Reliability Data

For the analysis of the data, a rating of .70 or higher will be considered a reliable measurement of the instrument. The reliability score is the intraclass correlation using Cronbach's Alpha described with the lower and upper bounds set to a 95% confidence interval. An ANOVA was run that calculated the grand mean of each group's assessment. The data was calculated using IBM SPSS 21 software.

Results of “I Like It”

Rater Groups	Reliability Score	Lower/Upper	Grand Mean
Teachers	.50	-.23-.88	3.07
Students	.71	.45-.88	3.42
Adults	.93	.79-.99	2.68

Results of “It was Creative”

Rater Groups	Reliability Score	Lower/Upper	Grand Mean
Teachers	.76	.40-.94	3.27
Students	.76 (.84)*	.69-.93	3.66
Adults	.88	.63-.99	2.78

*It should be noted one 6th grade student rater assessed every product as a 5, with the logic that “I thought everything was creative. All 12 products were creative because they were different from each other” (interview, 2013a). When removing that rater from the results, the reliability of 6th grade students assessing creativity dropped from .84 down to .76.

Results of “It was Unique”

Rater Groups	Reliability Score	Lower/Upper	Grand Mean
Teachers	.56	-.09-.89	3.20
Students	.67	.37-.87	3.63
Adults	.70	.07-.96	3.02

Results of “It Showed the Quality I Expect”

Rater Groups	Reliability Score	Lower/Upper	Grand Mean
Teachers	.85	.63-.96	3.44
Students	.67	.37-.87	3.63
Adults	.33	-1.07-.92	3.31

Interpretation of Reliability Data and Rater Responses

It was Creative is the only rating to show a reliable score throughout the rating process by every group. This would support the use of Amabile’s consensual assessment technique for teachers to rate creativity, but what does it say about the rest of the subjective judgments made by each rater group, and why is creativity being rated reliably while other subjective judgments

are not? These two questions also influenced why I asked the rating groups to also rate liking a product, the uniqueness of a product, and to rate the product for demonstrating the quality expected for this skill level.

For the teacher group, while they did not reliably assess if they liked a composition (.50) or found it unique (.56), there was a strong score for assessing the composition for quality expected of this skill level (.85). This would lead me to believe that teachers are rating creativity based on the skill level they expect. When asked why they rated a product as creative, teachers argued with each other for the reason behind their ratings (interview, 2013b). String teacher raters seemed to each be listening for different techniques being used, yet their results were reliable and consistent because they used their expertise as strings teachers to find specifics that made the product creative. Their differences about what made a product creative was inconsequential in regards to the fact that they all found different reasons for a product to be creative. This is an example of raters in this study using their personal schema as music teachers to expertly judge creativity. We may all view a product as creative for different reasons. If we're using our knowledge of quality expected based on the skill level of the composer we are rating, I would argue this to be a solid reasoning for assessing creativity and the validity of the consensual assessment technique being used.

We must then ask why the other groups are reliably assessing creativity using this model. I will start with the 6th grade students' review of their peers. Like the teachers, we may expect the students to be expert judges of quality (expected for their skill level). After all, these students' entire journey as musicians had up to this point been made together and they themselves all just participated in this project. However, their reliability score for *quality expected of this skill level* was just below a strong rating at .67. To explain this, I took in to

consideration students' reliability scores for *liking* a composition and finding a composition *unique* as I did with the 5th-6th grade teachers.

The 6th grade students had a reliability score of .71 when assessing *liking* it, compared to the relatively low score of .50 by the teachers. They also demonstrated a higher reliability for rating *uniqueness*, .67 for peer students to .56 by the teachers. When examining the non-musician adults group I found a high reliability to assess liking a product (.93), a good enough reliability to assess uniqueness (.70), but a low reliability to assess the quality expected for the skill level (.33). The two non-teacher groups are reliably rating creativity using the consensual assessment technique, but it would not initially appear they are doing so because they are expert judges of the quality expected of this skill level. For that reason, I dismiss those groups as expert judges regardless of their ability to reliably assess creativity. In looking at the correlations between the judgments being assessed, evidence is shown to support this claim and discussed below.

Interpretation of Correlations between Raters

It is important to distinguish why teachers are considered expert raters using the consensual assessment technique. If the other groups rate as reliable raters of creativity despite potentially not being expert judges, what does that tell us? Examining correlations between the raters and their ratings provided potential insight into the reasoning behind their assessments. A sampling of raters in each group was chosen at random and their correlations were calculated using IBM SPSS 21.

Student Correlations

Student 1		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.592* .043	.296 .043	.649* .023
Creative	Pearson Sig. (2-tailed)	.592* .043	1	.428 .165	.548 .065
Unique	Pearson Sig. (2-tailed)	.296 .351	.428 .165	1	.586* .045
Quality for Skill Level	Pearson Sig. (2-tailed)	.649* .023	.548 .065	.586* .045	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Student 1 has few strong correlations. As expected there was a moderate correlation between liking a product and rating creativity. In fact, the strongest factor for this rater rating a composition's creativity was how much they liked it. Quality expected for skill level showed a correlation between everything but creativity, which would suggest they are not an expert rater.

Student 2		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.605* .037	.753** .005	.568 .054
Creative	Pearson Sig. (2-tailed)	.605* .037	1	.540 .070	.478 .116
Unique	Pearson Sig. (2-tailed)	.753** .005	.540 .070	1	.469 .124
Quality for Skill Level	Pearson Sig. (2-tailed)	.568 .054	.478 .116	.469 .124	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Student 2 has correlations similar to student 1. Their strongest correlation between creativity was liking a product and the only significant correlation for rating creativity. If they liked a product, they also rated it as creative and unique. In this sample, there were no strong correlations between quality expected for the skill level and any other rating. This rater would not be considered an expert judge.

Student 3		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.574 .051	.581* .048	.735** .006
Creative	Pearson Sig. (2-tailed)	.574 .051	1	.790** .002	.593* .042
Unique	Pearson Sig. (2-tailed)	.581* .048	.790** .002	1	.632* .027
Quality for Skill Level	Pearson Sig. (2-tailed)	.735** .006	.593* .042	.632* .027	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Student 3 would appear to be a valid rater as liking a product was their lowest correlation for ranking something as creative. However, their strongest correlation for rating something creative is rating something unique. It would appear in this situation there is not much distinction between the two judgments. The fact that the lowest correlation for quality expected for this skill level is also creativity, therefore this rater is not considered an expert judge.

Student 4		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.069 .831	.617* .033	.680* .015
Creative	Pearson Sig. (2-tailed)	.069 .831	1	.580* .048	-.103 .749
Unique	Pearson Sig. (2-tailed)	.617* .033	.580* .048	1	.358 .253
Quality for Skill Level	Pearson Sig. (2-tailed)	.680* .015	-.103 .749	.358 .253	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

The final student rater in this sample continues to show the inconsistency of the student raters. In this case there was no correlation between liking a product and rating its creativity. The only correlation between rating creativity was if they found it unique. In this case, again we see the student not separating uniqueness from creativity with no correlation between quality expected for the skill level and creativity.

Across the student groups we see some correlations between liking a product or finding it unique and the creativity assessment. In no cases does the quality expected for this skill level

correlate with the ratings for creativity. I suggest that this would need to be a necessary correlation when assessing creativity. Without that correlation, it would appear judges are relying much more on the factor of whether they liked a product or found it unique, and those are not valid reasons for assessing creativity in the this study.

Non-Musician Adult Correlations

Adult 1		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.788** .002	.713** .009	.873** .000
Creative	Pearson Sig. (2-tailed)	.788** .002	1	.744** .005	.737** .006
Unique	Pearson Sig. (2-tailed)	.713** .009	.744** .005	1	.558 .060
Quality for Skill Level	Pearson Sig. (2-tailed)	.873** .000	.737** .006	.558 .060	1

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

The first non-musician adult rater showed strong correlations throughout the rating process. This is telling us they are considering multiple factors when assessing creativity and doing so consistently. Despite this positive correlation, this study shows that liking a product or finding it unique has a stronger correlation than the quality expected for skill level, and so this rater is not considered an expert judge.

Adult 2		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.944** .000	.674* .016	.710** .010
Creative	Pearson Sig. (2-tailed)	.944** .000	1	.700* .011	.745** .005
Unique	Pearson Sig. (2-tailed)	.674* .016	.700* .011	1	.614* .034
Quality for Skill Level	Pearson Sig. (2-tailed)	.710** .010	.745** .005	.614* .034	1

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

The second non-musician adult rater again shows strong correlations throughout each category. A positive correlation exists for quality expected for skill level as creativity has the

strongest correlation. However, this correlation is weaker between rating a product as liking it and finding it creative, and for that reason, this rater is not considered an expert judge in this study.

Adult 3		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.606* .037	.468 .125	.322 .307
Creative	Pearson Sig. (2-tailed)	.606* .037	1	.803** .002	.538 .071
Unique	Pearson Sig. (2-tailed)	.468 .125	.803** .002	1	.384 .217
Quality for Skill Level	Pearson Sig. (2-tailed)	.322 .307	.538 .071	.384 .217	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Our final sample from the non-musician adult group shows correlations very similar to the student group. No correlations showed between quality expected for the skill level. I saw a positive correlation between liking a product and finding it unique when assessing creativity. Since there is a missing correlation between quality expected for this skill level, I assert this judge is not an expert rater.

As with the student group, correlations between liking and finding a unique composition were stronger than the correlation between creativity and the quality expected for a student's skill level. This could explain why students (.76) and non-musician adults (.88) scored as reliable raters of creativity using the consensual assessment technique despite the fact that they are not considered expert judges in this study. The fact that they scored unreliable for expected quality (.67) for the students, (.33) for the non-musician adults, would support the low correlations between their ratings of creativity and expected skill quality and the fact that they are not considered expert judges. High reliability for liking a product and strong correlations between liking a composition and rating creativity would also support this idea. If we find that the teachers' correlation between expected quality and creativity is higher than the correlation

between assessing creativity and liking or the uniqueness of a composition, they will be considered expert judges in this study.

Teacher Correlations

Teacher 1		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.878** .000	.748** .005	.649* .023
Creative	Pearson Sig. (2-tailed)	.878** .000	1	.829** .001	.922** .000
Unique	Pearson Sig. (2-tailed)	.748** .005	.829** .001	1	.771** .003
Quality for Skill Level	Pearson Sig. (2-tailed)	.877** .000	.922** .000	.771** .003	1

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

The first teacher shows a strong correlation between all ratings. However, their strongest correlation is incredibly high and is quality expected for the skill level and creativity. This is to be expected from an expert judge.

Teacher 2		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.735** .006	.679* .015	.767* .004
Creative	Pearson Sig. (2-tailed)	.735** .006	1	.890** .000	.837** .001
Unique	Pearson Sig. (2-tailed)	.679* .015	.890** .000	1	.757** .004
Quality for Skill Level	Pearson Sig. (2-tailed)	.767** .004	.837** .001	.757** .004	1

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

Here we find another teacher with strong correlations throughout the categories. The strongest correlation between creativity was uniqueness in this sample. However, the strongest correlation for expected quality for skill level was creativity. This rater may have considered uniqueness similarly to creativity, and potentially more so than the other teacher judges. The fact that quality expected for skill level has the strongest correlation with creativity would make this teacher an expert judge.

Teacher 3		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.641* .025	.762** .004	.766** .004
Creative	Pearson Sig. (2-tailed)	.641* .025	1	.639* .025	.892** .000
Unique	Pearson Sig. (2-tailed)	.762** .004	.639* .025	1	.582* .047
Quality for Skill Level	Pearson Sig. (2-tailed)	.766** .004	.892** .000	.582* .047	1

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

The third teacher would also be considered an expert judge. The strongest correlation for creativity is quality expected for skill level and vice versa. Again there are strong correlations throughout, but the most important one appears to be the strongest.

Teacher 4		Like	Creative	Unique	Quality for Skill Level
Like	Pearson Sig. (2-tailed)	1	.741** .006	.857** .000	.857** .000
Creative	Pearson Sig. (2-tailed)	.741** .006	1	.765** .004	.803** .002
Unique	Pearson Sig. (2-tailed)	.857** .000	.765** .004	1	.782** .003
Quality for Skill Level	Pearson Sig. (2-tailed)	.857** .000	.803** .002	.782** .003	1

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

The last teacher sample is very similar to the second teacher sample. Evidence exists that implicates this last teacher is also an expert rater. The strongest correlation to creativity is quality expected for skill level. The strongest correlation for quality expected for skill level in this example is if they liked it, however the correlation was statistically strong with creativity as well (.803).

An expert judge should need to demonstrate a few correlations in their ratings for them to be deemed reliable using the consensual assessment technique. High reliability will likely be seen when assessing creativity as well as quality expected for the skill level. Strong correlations are helpful when found between the two, stronger than the correlations between creativity and

liking a composition or finding it unique. When taking these factors in to consideration, it becomes clear why the teacher group can be considered expert judges despite the fact that all groups' reliably assessed creativity. The groups each produced different, though linear, results for the grand mean of their ratings. Students rated too high and non-musician adults too low if we are to consider teachers the expert raters.

Summary

It is very important to distinguish why the 5th-6th grade teachers reliably assessed creativity. In this study, teachers proved to be reliable raters of skill and ability, and the students and non-musician adults did not. The teachers' ratings were valid and they were considered expert judges because of the correlations with creativity seen only in their particular group. The ability to identify overall quality based on skill level is a much more valid reason for reliably rating creativity than a student or adult rating it creative because they liked it or found it unique. These findings should give credence to the teacher as an expert judge, despite the fact that certain groups may reliably assess creativity using the consensual assessment technique.

Reflective Questions

- 1) Do you agree with the author's analysis of the data that a positive correlation between "skill quality" and "creative" are valid reasons to accept this technique as authentic? Are there other criteria that should be assessed that would support this assessment technique as authentic?
- 2) Some psychologists believe that children do not have the ability to be creative because a level of mastery must be accomplished in the designated field before an individual can begin producing creative products. How do you feel about this sentiment? What concepts or skills might these music educators have been focusing on to reliably assess creativity if creativity does not exist at the novice skill level?

3) The author suggests a highly-qualified educator should be considered an expert judge in their given field and therefore has the insight to reliably assess creativity. Would you feel comfortable assessing creativity in your own classroom? If you used a composition project as a summative assessment of understanding in music, would you choose to include creativity in the rubric? Why or why not?

Author Biography

Patrick Cooper is an Arizona native currently teaching general music, orchestra, music theatre, and music technology at a K-8 charter school in Chandler, Arizona. He serves on the board as secretary/treasurer for the Arizona Society for General Music and is also a double bass player in the Tempe Symphony Orchestra. Mr. Cooper earned his Bachelors and Masters Degrees from Arizona State University where he was awarded academic and performance-based scholarships, stipends, and grants. He hopes to continue his studies and begin a PhD program within the next few years, with the ultimate goal of being on a faculty at a major research institution.

Keywords: assessment, creativity, music, composition, children

Works Cited

Amabile, T.R. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5), p. 997-1013.

Cooper, P.K. (2013a, April 2). Personal interview.

Cooper, P.K. (2013b, April 3). Personal interview.

Cowdroy, R., Williams, A. (2006). Assessing creativity in the creative arts. *Art, Design & Communication in Higher Education*, 6(2), p. 97-117.

Hickey, M. (2001). An Application of Amabile's Consensual Assessment Technique for Rating Creativity of Children's Musical Compositions. *Journal of Research in Music Education*, 49(3), p. 234-244.

Sheridan, M., Byrne, C. (2002). Ebb and flow of assessment in music. *British Journal of Music Education*, 19, p. 135-143. doi:10.1017/S0265051702000220