

Cognition and the Assessment of Interaction Episodes in Jazz Improvisation

Brian C. Wesolowski
University of Georgia

The purpose of this study was to examine the cognitive processes related to interaction episodes in jazz improvisation through the development, testing, and validation of a rubric for assessing interaction in jazz improvisation. Four questions guided this study: (a) is a rubric an effective method for assessing performance achievement at undergraduate, graduate, and professional levels? (b) what is the interjudge reliability of a rubric used to assess interaction in jazz improvisation? (c) which of the rubric criteria maximally differentiate 3 performance achievement levels (i.e., undergraduate, graduate, and professional)? and (d) does a cognitive task ordering of difficulty exist in assessing jazz interaction? The rubric consisted of 3 melodic interaction criteria, 2 harmonic interaction criteria, and 3 rhythmic interaction criteria. A total of 55 expert judges evaluated 55 distinct jazz improvisation performances. The results indicated a statistically significant multivariate analysis of variance effect across all 3 performance achievement levels on each rubric item. The standardized discriminant function coefficients suggested that the 3 performance achievement levels were maximally differentiated by a canonical variate with greater weightings from all 3 melodic interaction criteria (i.e., “Reaction to implied [melodic] musical suggestions,” “Development of melodic musical ideas stemming from motivic interplay,” and “Adaptation to melodic interplay [call and response]”) and 1 rhythmic interaction criterion (i.e., “Coordination of rhythmical pulse [timing and synchronization]”). The results indicated a cognitive ordering of rater task difficulty: melodic interaction, rhythmic interaction, and harmonic interaction. A total of 66.33% of the variability was accounted for by performance achievement level, and interjudge reliability was estimated using Cronbach’s alpha ($\alpha = .91$).

Keywords: assessment, cognition, jazz, improvisation, interaction

Interactive episodes are often considered to be the foremost contributor of excitement and complexity in a jazz performance (Berliner, 1994, 1997; Hodson, 2007; Monson, 1996). According to Monson, “Good jazz improvisation is sociable and interactive just like a conversation; a good player communicates with the other players in the band. If this doesn’t happen, it’s not good jazz” (p. 84). A difficulty in evaluating and measuring interaction episodes in jazz improvisation lies in judging the ensembles’ overall performance achievement and creativity level (Collier, 1995). Research studies related to jazz improvisation creativity and performance achievement often exclude the variable of interaction due to the complexity of measuring performers’ cognitive and decision-making processes (Ciorbia, 2009; Limb & Braun, 2008; Madura, 1995; May, 2003; Norgaard, 2011; Smith, 2009; Ward-Steinman, 2008). Other studies acknowledge the collaborative nature of jazz improvisation but maintain that temporal cognition is evident only in real-time improvisational performance (Johnson-Laird, 2002; Mendonca & Wallace, 2004; Pressing, 1984). An additional problem with the evaluation of interaction in musical performance lies in the perception of observed behaviors. Interaction within an improvising jazz ensemble may not always be detected by observ-

able behaviors; in particular, the presence of subconscious cognitive and affective behaviors can potentially obscure the ability to perceive interaction by both the observer and the participants themselves (Seddon, 2005).

Further investigation into the cognitive processes of jazz interaction and performance achievement level is needed in order to broaden the scope of teaching and learning in jazz performance. The purpose of this study was to examine the cognitive processes related to interaction episodes in jazz improvisation through the development, testing, and validation of a rubric for assessing interaction in jazz performance. The research questions that guided this study include (a) is a rubric an effective method for assessing the performance achievement of interaction at undergraduate, graduate, and professional levels? (b) what is the interjudge reliability of a rubric used to assess interaction in jazz improvisation? (c) which of the rubric criteria maximally differentiate three performance achievement levels (i.e., undergraduate, graduate, and professional)? and (d) does a cognitive task ordering of rater difficulty exist in assessing jazz interaction?

Interaction, Communication, and Jazz Performance

According to Sawyer (1992), nonverbal interaction in jazz performance is derived from a balance of conscious and unconscious processes acting simultaneously. More specifically, interaction includes the conscious response to syntactical elements of musical knowledge (i.e., sympathetic attunement) and/or an unconscious collaborative aesthetic judgment stemming from ensemble accordance (i.e., empathetic attunement) (Seddon, 2005). According to Seddon:

Brian C. Wesolowski, Department of Music Education, Hugh Hodgson School of Music, University of Georgia.

Correspondence concerning this article should be addressed to Brian C. Wesolowski, Department of Music Education, Hugh Hodgson School of Music, University of Georgia, Athens, GA 30602-7287. E-mail: bwes@uga.edu

At one level this attunement is sympathetic, allowing for a cohesive performance revolving around sharing stocks of musical knowledge. At another level this attunement becomes empathetic, which is differentiated from sympathetic by collaborating musicians either consciously or subconsciously engaging in the psychological processes of decentring and introspection in order to become empathetically attuned. Once empathetically attuned, an atmosphere of trust allows for creative risk taking, which can result in the production of spontaneous musical utterances that may be regarded as examples of empathetic creativity (p. 58).

Empathetic attunement has also been described as ensemble groove or group flow (Berliner, 1994, 1997; Csikszentmihalyi, 1990, 1996; Kenny & Gellrich, 2002; Sawyer, 2006).

As Berliner (1994) explains, “. . . while attending to their own parts—assessing inventive material and selecting elements for development—performers must constantly exercise musical peripheral vision to make similar assessments about neighboring parts as they endeavor to predict their courses” (p. 364). Examples of such interaction include the subtle nuances and relationships between instruments in terms of synchronicity, reaction to the types of prevalent subdivisions being used (i.e., rhythmic density), adjustments to rhythmic patterns such as polyrhythms and on-beat/off-beat accents, adjustments to harmonic patterns of textural density and chord substitutions, reaction to the introduction of melodic fragments (i.e., melodic invention), and reaction to stylistic ideas engrained in the vocabulary of the jazz tradition (i.e., quotations, recurring vocabulary patterns, logical phrase construction, musical suggestion, and musical expectation) (pp. 348–379). Furthermore, Rinzler (1986) provides examples of various types of interaction that may occur specifically between a soloist and rhythm section: (a) call and response; (b) use of fills; (c) accenting the ends of formal units; (d) the harmonic or rhythmic repetition of a motive (common motive); and (e) response to the “peaks” of the soloist (pp. 156–157).

In addition to the “stocks of musical knowledge” that Berliner (1994) and Rinzler (1986) describe, a performer’s degree of creativity is a factor in the level of attunement achieved (Sawyer, 2003). According to Rinzler, a performer may engage in three major degrees of creativity related to interaction: (a) the performers may produce music . . . with little or no creative impulse; (b) performers may produce music by creatively fulfilling their own individual musical function but not interacting with fellow musicians; and (c) performers may produce music by creatively interacting with the other musicians in the group (p. 156).

To assess interaction in a jazz performance from a pedagogical approach, one must consider several dimensions related to the musical development of the ensemble: (a) degree of obtained vocabulary (i.e., What subgenres of jazz are the students knowledgeable of? What standardized vocabulary have the students acquired?); (b) vocabulary development (i.e., Are the students able to connect multiple ideas throughout chord progressions? Are the students still developing their vocabulary over specific chord progressions?); (c) aural ability (i.e., Are the students able to recognize musical patterns and replicate them? Are the students able to interpret a musical idea? Are the students able to develop musical ideas from musical motives occurring within the ensemble?); and (d) level of response (i.e., Can the students listen and improvise simultaneously? To what degree can the students hear a phrase, internalize it, replicate it, apply it, and adapt it? How critically and in depth can the students hear? Can the student listen across the

ensemble to multiple instrumentalists at one time? Does listening disrupt the flow of their performance?) By relating these dimensions to interaction episodes via a multidimensional rubric, the pedagogy and assessment of more elusive aspects of jazz improvisation may be enhanced. In addition, the analysis of such an assessment tool may provide further insight into what central factors contribute to the cognition and perception of interaction episodes at multiple achievement levels.

Jazz Pedagogy, Rubrics, and Music Performance Assessment

A frequent criticism of academic-based jazz improvisation pedagogy is the emphasis on syntactical technique-driven development at the sacrifice of creative improvisational thinking (Prouty, 2002). According to Prouty (2012), “In evaluating student performers, at least two main forces are at work, one of which is under control of jazz educators (the teaching of technique, e.g.), while the other largely is not (individual creativity)” (p. 68). Institutional demands have led educators to codify pedagogical methods that are easily measurable and quantified, such as cognitive ability-type testing of syntax and vocabulary development (Nicholson, 2005). This allows students the opportunity to construct, replicate, and perform an improvised solo using acquired syntax and vocabulary alone while demonstrating little creativity (Prouty, 2008). The subjective nature of assessing creative elements of jazz improvisation, such as ensemble interaction and communication, is proven difficult to assess in an objective concrete manner.

Current trends in higher education assessment are moving away from traditional cognitive-type tests and shifting more toward authentic assessments (Dochy, Gijbels, & Segers, 2006). Authentic assessments reflect more valid data about student competence through their engagement with meaningful tasks that demonstrate essential skills and knowledge (Darling-Hammond, & Snyder, 2000). In addition, authentic assessments aim at assessing higher order thinking processes and competencies as opposed to the factual knowledge of cognitive testing. This, in turn, has led to a strong interest in various types of performance assessments that capture more elusive aspects of learning (Jonsson & Svingby, 2007).

Rubrics offer a mechanism for assessing higher-order thinking by documenting students’ developmental progress as well as providing clear levels of achievement to which they may aspire. A rubric is a form of a criteria-specific performance scale that divides a task into constituent parts and offers detailed descriptions of the performance levels for each part (Wesolowski, 2012). Research has demonstrated many benefits of the implementation of rubrics to the educational process. Research reports have shown evidence that rubrics (a) provide diagnostic feedback that can identify areas for improvement in instruction (Song, 2006); (b) serve instructional as well as evaluative purposes (Andrade & Du, 2005); (c) support teaching and student learning (Andrade & Du, 2005; Osana & Seymour, 2004; Reitmeier, Svendsen, & Vrchota, 2004; Schneider, 2006); and (d) reflect higher achievement and deeper learning by students who have rubrics to guide their work (Petkov & Petkova, 2006; Reitmeier et al., 2004). Student perception has been shown to reflect positively for their implementation. Graduate and undergraduate students value rubrics because they “clarify the targets for their work, allow them to regulate their progress, and make grades or marks transparent and fair” (Reddy & An-

Criteria	Beginning	Developing	Proficient	Accomplished
Melodic				
1. Development of melodic musical ideas stemming from motivic interplay	<input type="checkbox"/> 1 pt. Ensemble seldom develops melodic ideas and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes develops melodic ideas and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually develops melodic ideas and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble develops melodic ideas with artistic awareness and interaction never disrupts improvisational fluency
2. Reaction to implied (melodic) musical suggestions	<input type="checkbox"/> 1 pt. Ensemble seldom develops implied musical suggestions and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes develops implied musical suggestions and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually develops implied musical suggestions and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble develops implied musical suggestions with artistic awareness and interaction never disrupts improvisational fluency
3. Adaptation to melodic interplay (call and response)	<input type="checkbox"/> 1 pt. Ensemble seldom reacts to melodic figurations and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes reacts to melodic figurations and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually reacts to melodic figurations and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble reacts to melodic figurations with artistic awareness and interaction never disrupts improvisational fluency
Harmonic				
4. Development of harmonic ideas stemming from harmonic interplay	<input type="checkbox"/> 1 pt. Ensemble seldom develops harmonic ideas and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes develops harmonic ideas and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually develops harmonic ideas and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble develops harmonic ideas with artistic awareness and interaction never disrupts improvisational fluency
5. Adaptation to harmonic interplay	<input type="checkbox"/> 1 pt. Ensemble seldom reacts to harmonic substitutions and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes reacts to harmonic substitutions and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually reacts to harmonic substitutions and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble reacts to harmonic substitutions with artistic awareness and interaction never disrupts improvisational fluency

Figure 1. Analytic rubric for assessing interaction in jazz improvisation.

drade, 2010). Instructors found that rubrics provide an objective basis for evaluation, more consistent, reliable, and efficient grades (Campbell, 2005), and facilitates a change in evaluation procedures from “subjective observations to specific performances” (Reitmeier et al., 2004, p. 18).

A primary difficulty with music performance assessment is managing its subjective nature (Radocy, 1986). Rubrics can be used to overcome its subjective nature and to help develop a set of

guidelines for clearly assessing student work and thus to increase objectivity in the evaluation of performances (Wesolowski, 2012). In the field of music education, rubrics have been developed and empirically tested for the influence of certain variables on solo and small ensemble festival ratings (Bergee & Platt, 2003), effects of imagery-based instruction on musical expression (Woody, 2006), the relationships of keyboard ownership to performance ratings (Price, 2007), choral festival adjudication (Norris & Borst, 2007), undergraduate performance juries (Ciorba & Smith, 2009), and collegiate-applied studio instruction (Parkes, 2005). The implementation of statistically reliable and valid rubrics can serve as a valuable assessment tool for music educators to facilitate learning and improve teaching in the classroom as well as aid in the assessment of complex behaviors.

Method

Measure Development

The proposed measure for the assessment of interaction in jazz performance was a 4-point analytic rubric, using eight dimensions of evaluative criteria (see Figure 1). The levels were gleaned from Rinzler’s (1986) outline of three major degrees of creativity and Berliner’s (1994) description of how jazz ensemble members engage in nonverbal interaction during a performance. The language used in rubrics is considered to be one of the most important

Criteria	Beginning	Developing	Proficient	Accomplished
Rhythmic				
6. Coordination of rhythmic pulse (timing and synchronization)	<input type="checkbox"/> 1 pt. Ensemble seldom reacts to changes in timing and synchronization and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes reacts to changes in timing and synchronization and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually reacts to changes in timing and synchronization and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble always reacts to changes in timing and synchronization with artistic awareness and interaction never disrupts improvisational fluency
7. Adaptation to rhythmic interplay (call and response)	<input type="checkbox"/> 1 pt. Ensemble seldom reacts to rhythmic figurations and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes reacts to rhythmic figurations and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually reacts to rhythmic figurations and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble reacts to rhythmic figurations with artistic awareness and interaction never disrupts improvisational fluency
8. Adaptation to rhythmic density	<input type="checkbox"/> 1 pt. Ensemble seldom reacts to changes in rhythmic density and interaction extremely disrupts improvisational fluency	<input type="checkbox"/> 2 pts. Ensemble sometimes reacts to changes in rhythmic density and interaction moderately disrupts improvisational fluency	<input type="checkbox"/> 3 pts. Ensemble usually reacts to changes in rhythmic density and interaction slightly disrupts improvisational fluency	<input type="checkbox"/> 4 pts. Ensemble reacts to changes in rhythmic density with artistic awareness and interaction never disrupts improvisational fluency

Figure 1. (continued)

factors for achieving reliable and valid large-scale assessments (Moni, Beswick, & Moni, 2005; Tierney & Simon, 2004). The complexity of this rubric is a consequence of the multifaceted nature of aurally examining interaction in jazz performances as well as the intricate process of engaging in interaction episodes in improvisational performance. The rubric was constructed to accommodate the population and advanced level of students being assessed (Green & Bowser, 2006). The rubric was piloted using freshman and sophomore undergraduate jazz major improvisational performances ($N = 10$), graduate jazz major improvisational performances ($N = 10$), and professional improvisation performances ($N = 10$). Undergraduate, graduate, and faculty members at three universities with accredited jazz programs provided the recordings. The selected recordings for the pilot study demonstrated a wide range of performance ability. The judging panel ($N = 5$) consisted of full-time jazz studies faculty members from varying universities who possess expertise in jazz performance and teaching. The panel was instructed to listen to the performances as many times as needed and to judge the ensemble using the specified measurement tool. Initial data were analyzed and feedback regarding the measurement tool was solicited from the panel. The rubric was revised and edited based on suggestions for clarity, practicality, redundancy, and continuity.

Subjects and Judges

Expert judges ($N = 55$) were selected based on their jazz education experience, jazz performance experience, academic background, and availability. The volunteer judges were drawn from a pool of professional musicians adept in the jazz idiom and university professors holding full-time positions in jazz studies departments. The judges represented diverse educational backgrounds, teaching expertise, and performing experiences. Before the evaluation, each judge viewed an instructional packet that outlined the premise of the research study and clearly defined the task expectations and degrees of proficiencies of the measurement instrument. The judges and recordings used in the pilot study were not used in the full-scale study.

Following the methodology of Smith (2009), each judge was provided two anchor examples representing a beginning level and an accomplished level of interaction in an improvisational performance. The anchor examples were offered to provide a reference point for judging criteria that spanned the range of the provided measurement tool. Each judge was then given one anonymous recording. The recordings were anonymous to control for judges' biases for performance level. The instrumentation of each recording was a jazz quartet, including a brass or woodwind soloist, piano, bass, and drums. Each recording was performed in a swing style using functional harmony with a standard song form or blues form as the improvisational vehicle. The recordings were randomly selected from a pool of noncommercially released undergraduate, graduate, and professional improvisational performances. The recorded performances averaged between 45 s and 1 min and 30 s. Each judge was instructed to evaluate the ensemble using the provided measurement tool within the range of skill illustrated by the two provided anchor recordings.

An a priori power analysis was conducted to estimate a sufficient sample size that would yield a medium to large effect size (i.e., .50–.80) with a power of .80 and alpha level of .05. The

projected sample size ranged from 18 to 24 evaluations. A total of 55 expert judges evaluated 55 distinct jazz improvisation performances (i.e., undergraduate, $n = 17$; graduate, $n = 20$; professional, $n = 18$). This resulted in a 73% response rate.

Results

A one-way multivariate analysis of variance (MANOVA) was conducted to test the hypothesis that there would be significant mean differences between performance levels (i.e., undergraduate, graduate, and professional) and the eight criteria comprising the rubric. Before conducting the MANOVA, several assumptions were checked. Box's M value was interpreted as nonsignificant (104.93, $p = .32$), indicating that the covariance matrices were not significantly different; thus, the assumption of homoscedasticity was maintained (Huberty & Petroskey, 2000). The Shapiro–Wilk test of normality indicated that the population was distributed normally ($p > .05$) for each of the achievement levels. A series of Pearson correlations were performed between all of the dependent variables. Table 1 indicates that they were moderately correlated, ranging from .33 to .76 (Meyer, Gampst, & Guarino, 2006). Additionally, an evaluation of the Mahalanobis distances ($\alpha = .001$) indicated no significant outliers, and a series of Levene's F tests satisfied the homogeneity of variance assumption. Fifty-five evaluators independently judged 55 distinct recordings; therefore, independence was assumed.

Table 2 presents the descriptive statistics for each rubric criterion and sum scores presented in this research study. A statistically significant MANOVA effect was obtained ($\lambda = .33$, $F_{(16,90)} = 4.10$, $p < .001$, $\eta^2 = .42$). The effect size .42 indicates that 42% of the variance was accounted for by performance achievement level. A series of one-way analysis of variance (ANOVA) on each of the eight dependent variables was conducted as a follow-up test to the MANOVA. All of the ANOVAs were statistically significant (See Table 3).

A series of Bonferroni-adjusted post hoc analyses were performed to examine individual mean difference comparisons across the three performance achievement levels and all eight rubric criteria. As seen in Table 4, the results revealed that all post hoc mean comparisons were statistically significant.

A canonical discriminant function analysis was conducted to test the hypothesis that some rubric criteria may maximally differentiate the three performance levels. Two eigenvalues and canonical variables were extracted; however, only the first extraction indicated significance. The first eigenvalue was equal to 1.73 and accounted for

Table 1
Pearson Correlations Associated With Ensemble Performance Ratings of the Rubric Criteria

Rubric criteria	1	2	3	4	5	6	7
Criterion 1	—						
Criterion 2	.68	—					
Criterion 3	.55	.34	—				
Criterion 4	.67	.52	.53	—			
Criterion 5	.51	.33	.50	.60	—		
Criterion 6	.43	.38	.47	.65	.48	—	
Criterion 7	.71	.57	.60	.73	.59	.53	—
Criterion 8	.76	.66	.56	.75	.54	.55	.70

Table 2
Descriptive Statistics for Rubric Criteria ($N = 55$)

Rubric criteria	Mean	SD	Variance
Criterion 1	2.75	.87	.79
Criterion 2	2.67	.90	.82
Criterion 3	2.98	.76	.57
Criterion 4	3.22	.79	.62
Criterion 5	3.20	.65	.42
Criterion 6	3.15	.73	.53
Criterion 7	3.09	.78	.60
Criterion 8	3.02	.78	.61

40% of the model variance (Wilks $\lambda = .33$, $F_{(16,90)} = 4.10$, $p < .001$). As seen in Table 5, the standardized canonical discriminant function coefficients demonstrate that the three performance levels were maximally differentiated by one canonical variate with the greatest weight from all three melodic interaction criteria ("Reaction to implied [melodic] musical suggestions," [.51] "Development of melodic musical ideas stemming from motivic interplay," [.31] "Adaptation to melodic interplay" [.32]) and one rhythmic interaction criterion ("Coordination of rhythmical pulse" [.50]). Therefore, the results indicated a cognitive ordering of rater task difficulty: melodic interaction, rhythmic interaction, and harmonic interaction. Canonically derived group means were then estimated for the three performance levels. As hypothesized, the professional group was associated with the highest mean ($M = 6.57$, $SD = .88$), the graduate group was associated with the second highest mean ($M = 5.92$, $SD = .92$), and the undergraduate group was associated with the smallest mean ($M = 4.08$, $SD = 1.17$).

Interjudge reliability was calculated using Cronbach's alpha lower bound estimate of reliability. The result for all raters ($N = 55$) was .91. The high alpha reliabilities provide an indication that there was a significant level of internal consistency among the judges' scores and the variability of individual criteria within the rubric. A total of 66.33% of the variability was accounted for by performance achievement level.

Multicollinearity was assessed using variance inflation factor values. All values were under the recommended value of 10, indicating that the collinear variables (i.e., item responses) do not demonstrate redundancy with the dependent variables (i.e., rubric items) (Hair, Anderson, Tatham, & Black, 1995).

Table 4
Post Hoc Mean Comparisons (Bonferroni Adjusted)

Performance level	N	M	SD
Undergraduate	17	4.65	.90
Graduate	2	6.60	1.00
Professional	18	7.49	.84
Total	55	6.29	1.47

Note. $F_{(2,55)} = 43.44$; $p < .001$.

Discussion

The purpose of this study was to examine cognitive processes related to interaction episodes in jazz improvisation through the generation, testing, and validation of a rubric based on interaction episodes in jazz improvisation. The results of this study provide evidence that the prescribed rubric can be used effectively to evaluate interaction in jazz improvisation with significant reliability ($\alpha = .91$). Specifically, the rubric was able to discriminate between undergraduate, graduate, and professional achievement levels across each criterion outlined in the rubric. As hypothesized, the overall means across the three performance levels indicated that improvisers perform with greater attention to ensemble interaction and with greater improvisatory fluidity at higher performance levels. This indicated that the rubric has a high level of content validity. Strong content validity along with the moderate to strong correlations between the rubric criteria and significant reliabilities supports Fiske's (1983) argument that subjective performance-based activities can be effectively assessed in a holistic and objective manner.

The standardized discriminant function coefficients suggested that the three performance levels were maximally differentiated by a canonical variate with greater weightings from three items from the melodic criteria (i.e., "Reaction to implied (melodic) musical suggestions," "Development of melodic musical ideas stemming from motivic interplay," "Adaptation to melodic interplay [call and response]"), one item from the rhythmic criteria (i.e., "Coordination of rhythmical pulse [timing and synchronization]"), and no items from the harmonic criteria. This implied a cognitive ordering of task difficulty with respect to the canonical variate: melodic interaction ($n = 3$), rhythmic interaction ($n = 1$), and harmonic interaction ($n = 0$). Moreover, this sheds some light on the

Table 3
One-Way ANOVAs With Rubric Criteria as Dependent Variables and Performance Level as Independent Variable

Rubric criteria	$F_{(16,90)}$	Undergraduate		Graduate		Professional	
		M	SD	M	SD	M	SD
Criterion 1	22.57	1.88	.60	2.95	.76	3.33	.59
Criterion 2	19.18	1.82	.53	2.90	.72	3.22	.81
Criterion 3	11.44	2.41	.62	3.05	.69	3.44	.62
Criterion 4	13.58	2.59	.87	3.30	.57	3.72	.46
Criterion 5	6.84	2.76	.56	3.35	.67	3.44	.51
Criterion 6	14.16	2.59	.62	3.15	.67	3.67	.49
Criterion 7	15.66	2.41	.62	3.25	.64	3.56	.62
Criterion 8	17.48	2.29	.69	3.25	.55	3.44	.62

Note. $N = 55$; $p < .001$.

Table 5
Discriminant Function Coefficients Associated With the MANOVA

Rubric criteria	Raw	Standard	Structure
Criterion 1	.66	.43	.71
Criterion 2	.72	.51	.65
Criterion 3	.50	.32	.50
Criterion 4	.09	.06	.55
Criterion 5	.01	.01	.38
Criterion 6	.83	.50	.55
Criterion 7	.08	.04	.59
Criterion 8	.07	.11	.61

Note. $N = 55$; eigenvalue was equal to 1.73 with a canonical correlation of .80. Raw = unstandardized coefficients; standard = standardized coefficients; structure = structure coefficients.

cognitive processes related to interaction in jazz improvisation at various achievement levels. These four items carried the most weight in the canonically derived group means. Although all items in the rubric were able to discriminate between the three performance achievement levels, the four highlighted items indicate where professional, graduate, and undergraduate soloists demonstrate the greatest differences. Therefore, one can speculate that these particular items may necessitate the most amounts of musical ability, adaptation, responsiveness, critical thinking, and/or creativity. A strong aural capacity is needed to engage in interaction at the melodic level. Interestingly, item number two (Reaction to implied [melodic] musical suggestions) carried the greatest weight in the canonical variate (.51). This item not only calls for a strong aural capacity, but also a strong foundation in jazz vocabulary, knowledge of repertoire, and stylistic convention. The only rhythmic item underscoring the canonical correlate was “Coordination of rhythmical pulse (timing and synchronization).” This strengthens Seddon’s (2005) notion of empathetic attunement, where advanced levels of interaction move beyond vocabulary-based communication and into groove-oriented communication and accordance. Here, this is sense of “togetherness” can be evidenced in ensemble synchronization.

The suggested cognitive-task ordering provides a great amount of diagnostic and analytical value that may inform pedagogy and improve students’ awareness in the aural and interactive aspects of jazz improvisation. Pedagogically, it may provide a framework for expectations of student improvisers. Undergraduates who strive to play at a professional level can focus on the four items with the greatest canonical weightings to increase their perceived ability to interact in the ensemble setting. A developed aural ability will allow students to melodically interact with the ensemble, and increased attention to ensemble synchronization (i.e., ensemble accordance) can improve the perception of interaction in the ensemble.

Bringing to light the construct validity of the measurement tool, one can only speculate on the relationship of listener perception to acoustical occurrence within the limitations of this study. The rubric identifies the cognitive rater task difficulty: melodic interaction, rhythmic interaction, and harmonic interaction. Arguably, harmonic interaction may not have been perceived, thus causing its subjacent position in the cognitive task ordering. Additionally, melodic interaction may be the most perceptually apparent form of interaction in jazz improvisation. Listeners’ focus may be directed to melodic interplay due to the aural/oral nature and tradition of jazz

stemming from the historical, pedagogical, and musicological significance of call and response. Future experimental studies that examine the connection between these variables and listeners’ perception of interaction could potentially shed light on these perceptual and acoustical relationships. Furthermore, the perceived relationships between improvisation quality, interaction quality, and rater score are unclear. Raters may have potentially evaluated quality of the performance in lieu of occurring interactions. Experimental research that parses apart these relationships may lead to a clearer connection and understanding of the perception of ensemble interaction.

The development of valid and reliable assessment tools cannot only improve the teaching and learning processes, but can also provide insight into the more elusive aspects of musical behavior. Jazz improvisation and ensemble interaction are multifaceted and complex musical behaviors that use higher order thinking and creative competencies. The newly constructed assessment tool provided in this study may aid in the process of recognizing, diagnosing, and communicating new methods for improvement in an individual student’s jazz improvisatory development, solo construction, and execution.

References

- Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation, 10*, 1–11.
- Bergee, M. J., & Platt, M. C. (2003). Influence of selected variables on solo and small-ensemble festival ratings. *Journal of Research in Music Education, 51*, 342–353. doi:10.2307/3345660
- Berliner, P. (1994). *Thinking in jazz: The infinite art of improvisation*. Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226044521.001.0001
- Berliner, P. (1997). Give and take: The collective conversation of jazz performance. In R. K. Sawyer (Ed.), *Creativity in performance* (pp. 9–41). Greenwich, CT: Ablex Publishers.
- Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgment is involved: The features of an e-marking tool. *Assessment and Evaluation in Higher Education, 30*, 529–537. doi:10.1080/02602930500187055
- Ciorbia, C. R. (2009). Predicting jazz improvisation achievement through the creation of a path analytical model. *Bulletin of the Council for Research in Music Education, 180*, 43–57.
- Ciorba, C. R., & Smith, N. Y. (2009). Measurement of instrumental and vocal undergraduate performance juries using a multidimensional assessment rubric. *Journal of Research in Music Education, 57*, 5–15. doi:10.1177/0022429409333405
- Collier, G. (1995). *Interaction: Opening up the jazz ensemble*. Berlin: Advance Music.
- Csikszentmihalyi, M. (1990). The domain of creativity. In M. A. Runco and R. S. Albert (Eds.), *Theories of creativity* (pp. 190–212). Newbury Park, CA: Sage.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology for the third millennium*. New York: Harper Collins.
- Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education, 16*, 523–545. doi:10.1016/S0742-051X(00)00015-9
- Dochy, F., Gijbels, D., & Segers, M. (2006). Learning and the emerging new assessment culture. In L. Verschaffel, F. Dochy, M. Boekaerts, and S. Vosniadou (Eds.), *Instructional psychology: Past, present, and future trends*. Oxford, Amsterdam: Elsevier.
- Fiske, H. E. (1983). Judging musical performances: Method or madness? *Update: The Applications of Research in Music Education, 1*, 7–10.

- Green, R., & Bowser, M. (2006). Observations from the field: Sharing a literature review rubric. *Journal of Library Administration*, 45, 185–202. doi:10.1300/J111v45n01_10
- Hair, J. F., Anderson, J. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (5th ed.). New York: MacMillan.
- Hodson, R. (2007). *Interaction, improvisation, and interplay in jazz*. New York: Routledge.
- Huberty, C. J., & Petoskey, M. D. (2000). Multivariate analysis of variance and covariance. In H. Tinsley and S. Brown (Eds.) *Handbook of applied multivariate statistics and mathematical modeling*. New York: Academic Press. doi:10.1016/B978-012691360-6/50008-2
- Johnson-Laird, P. (2002). How jazz musicians improvise. *Music Perception*, 19, 415–442.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144. doi:10.1016/j.edurev.2007.05.002
- Kenny, B. J., & Gellrich, M. (2002). Improvisation. In R. Parncutt & G. E. McPherson (Eds.), *The science and psychology of music performance* (pp. 117–134). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195138108.003.0008
- Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS One*, 3, e1679.
- Madura, P. D. (1995). An exploratory investigation of the assessment of vocal jazz improvisation. *Psychology of Music*, 23, 48–62. doi:10.1177/0305735695231004
- May, L. F. (2003). Factors and abilities influencing achievement in instrumental jazz improvisation. *Journal of Research in Music Education*, 51, 245–258. doi:10.2307/3345377
- Mendonca, D., & Wallace, W. A. (2004). *Cognition in jazz improvisation: An exploratory study*. Paper presented at the Cognitive Science Society Annual Meeting, Chicago, IL.
- Meyer, L. S., Gamst, G., & Guarino, A. (2006). *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: Sage Publishers.
- Moni, R. W., Beswick, E., & Moni, K. B. (2005). Using student feedback to construct an assessment rubric for a concept map in physiology. *Advances in Physiology Education*, 29, 197–203. doi:10.1152/advan.00066.2004
- Monson, I. T. (1996). *Saying something: Jazz improvisation and interaction*. Chicago, IL: University of Chicago Press.
- Nicholson, S. (2005). *Is jazz dead (or has it moved to a new address?)*. New York: Taylor and Francis.
- Norgaard, M. (2011). Descriptions of improvisational thinking by artist-level jazz musicians. *Journal of Research in Music Education*, 59, 109–127. doi:10.1177/0022429411405669
- Norris, C. E., & Borst, J. E. (2007). An examination of the reliabilities of two choral festival adjudication forms. *Journal of Research in Music Education*, 55, 237–251. doi:10.1177/002242940705500305
- Osana, H. P., & Seymour, J. R. (2004). Critical thinking in preservice teachers: A rubric for evaluating argumentation and statistical reasoning. *Educational Research and Evaluation*, 10, 473–498. doi:10.1080/13803610512331383529
- Parkes, K. A. (2005). *The effect of performance grading rubrics on the college music applied studio* (Doctoral dissertation). University of Miami, Coral Gables, FL.
- Petkov, D., & Petkova, O. (2006). Development of scoring rubrics for IS projects as an assessment tool. *Issues in Informing Science and Information Technology*, 3, 499–510.
- Pressing, J. (1984). Cognitive processes in improvisation. In R. Crozier & A. Chapman (Eds.), *Cognitive processes in the perception of art*. Amsterdam: North-Holland. doi:10.1016/S0166-4115(08)62358-4
- Price, H. E. (2007). Effect of keyboard ownership on keyboard performance in a music fundamental course. *International Journal of Music Education*, 25, 49–53. doi:10.1177/0255761407074891
- Prouty, K. (2002). *From Storyville to state university: The intersection of academic and nonacademic learning cultures in post-secondary jazz education* (Doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.
- Prouty, K. (2008). The “finite” art of improvisation: Pedagogy and power in jazz education. *Critical Studies in Improvisation*, 4, 1–15.
- Prouty, K. (2012). *Knowing jazz: Community, pedagogy, and canon in the information age*. Jackson: University Press of Mississippi.
- Radocy, R. E. (1986). On quantifying the uncountable in musical behavior. *Bulletin of the Council for Research in Music Education*, 88, 22–31.
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education*, 35, 435–448. doi:10.1080/02602930902862859
- Reitmeier, C. A., Svendsen, L. K., & Vrchota, D. A. (2004). Improving oral communication skills of students in food science courses. *Journal of Food Science Education*, 3, 15–20. doi:10.1111/j.1541-4329.2004.tb00036.x
- Rinzler, P. (1986). Preliminary thoughts on analyzing musical interaction among jazz performers. *Annual Review of Jazz Studies*, 4, 153–160.
- Sawyer, R. K. (1992). Improvisational creativity: An analysis of jazz performance. *Creativity Research Journal*, 5, 253–263. doi:10.1080/10400419209534439
- Sawyer, R. K. (2003). *Group creativity: Music, theatre, collaboration*. Mahwah, NJ: Erlbaum Assoc.
- Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation* (2nd ed.). New York: Oxford University Press, Inc.
- Schneider, J. F. (2006). Rubrics for teacher education in community college. *The Community College Enterprise*, 12, 39–55.
- Seddon, F. A. (2005). Modes of communication during jazz improvisation. *British Journal of Music Education*, 22(1), 47–61. doi:10.1017/S0265051704005984
- Smith, D. T. (2009). Development and validation of a rating scale for wind jazz improvisation performance. *Journal of Research in Music Education*, 56, 5–17.
- Song, K. H. (2006). A conceptual model of assessing teaching performance and intellectual development of teacher candidates: A pilot study in the US. *Teaching in Higher Education*, 11, 175–190. doi:10.1080/13562510500527701
- Tierney, R., & Simon, M. (2004). What’s still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research and Evaluation*, 9. Retrieved May 12, 2012 from <http://PAREonline.net/getvn.asp?v=9&n=2>.
- Ward-Steinman, P. M. (2008). Vocal improvisation and creative thinking by Australian and American university singers: A factor analytic study. *Journal of Research in Music Education*, 56, 5–17. doi:10.1177/0022429408322458
- Wesolowski, B. C. (2012). Understanding and developing rubrics for music performance assessment. *Music Educators Journal*, 98, 36–42. doi:10.1177/0027432111432524
- Woody, R. H. (2006). Musicians’ cognitive processing of imagery-based instructions for expressive performance. *Journal of Research in Music Education*, 54, 125–137.

Received April 30, 2013

Revision received September 30, 2013

Accepted October 10, 2013 ■