

Can expert listeners hear if a piece is improvised or composed?

*Andreas C. Lehmann & Reinhard Kopiez**

University of Music Würzburg, Germany

*Hanover University of Music and Drama, Germany

ABSTRACT

Although a performer knows if he or she is improvising or performing a rehearsed piece of music, this distinction may not be evident to the audience. However, it is possible that typical aspects of performance resulting from more deliberation and preparation or spontaneity induce perceptually salient cues that the listeners can utilize to infer whether a piece is composed or not. In other words, the listeners may be able to identify the music's generative type. The goal of our study was to test expert listeners' ability to discern between same-style compositions and non-compositions. Three composed pieces of different styles of music (classical, romantic, free jazz) were matched with highly similar non-composed excerpts. In a listening experiment, 102 college music students rated the excerpts on the scales "rehearsed/coherent" (RC) and "spontaneous/ improvised" (SI). Retest-reliabilities for the ratings were satisfactory. As expected, SI and RC ratings were negatively correlated (-.71). Using repeated measures ANOVA, significant main effects were found for style, generative type, and style by type interaction. Both, style effect and the interaction were elicited by the jazz piece. The listener rated its improvised example as more SI than the composed excerpt, which they rated more RC. Apparently, expert listeners had problems distinguishing improvised from same-style composed music of the classical and romantic era. In the case of the free jazz example they might have extracted performative cues (e.g., varying degrees of tightness in ensemble playing and rhythmic accuracy) that led them to correctly infer the generative type of the piece.

1. GENERATIVE TYPES: IMPROVISATION AND COMPOSITION

Musical practitioners as well as the musicological and educational literature differentiate between improvisation and composition, which we tend to call **generative types**, because they denote outcomes of generative musical processes. The constraints under which the improvisation and composition processes take place are different and typified. Commonly mentioned differences are four characteristics (time constraint, complexity, notation, coherence). These may, however, be less helpful for our understanding of musical creativity than is generally assumed (cf. Nettl, 1998; Pressing, 1998; see Lehmann, 2005, for a comprehensive review on generative processes). A clear theoretical distinction between these generative types (improvisation vs. composition) is difficult to draw because both activities and resulting "handicrafts" fulfill similar aesthetic functions. Also, not addressed here at all, the perception of improvisation may be guided by conventionalized

constructions between performers and audiences (Auslander, in press). But is it possible for the audience to discern one from the other perceptually?

1.1 Anecdotal evidence regarding the problem of discerning improvisation from composition

Reports by individual musicians suggest that they practice and memorize parts of their improvisations (cf. Sidran, 1995). Nettl (1998) suggests that to the uninitiated listener of a particular ethnic music, different-sounding renditions of a piece may still be intended to be the same piece; conversely, same-sounding renditions could indeed be intended as different-sounding. The 19th century German music critic Adolph Bernhard Marx was once visited by the famous piano virtuoso Friedrich Kalkbrenner, who improvised for about 15 minutes at the piano. Marx was impressed by Kalkbrenner's performance. A few days later, however, his admiration turned into anger when he received a package from a publisher with new compositions, among them Kalkbrenner's improvisation from few days before. Apparently, the work had already been composed and printed, and the critic was deceived by the boisterous pianist (Gerig, 1974, cit. in Lehmann & Ericsson, 1998, p. 76). Another interesting case is that of composer Giacinto Scelsi. He tape-recorded and subsequently edited his transcribed instrumental improvisations, thereby causing controversy among critics who even considered his method compositional fraud (Fox & Osmond-Smith, 2009). Many more examples could be given (see Lehmann & Kopiez, in press, for further details). Obviously, listeners have difficulties discerning the generative types based on mere listening.

1.2 Listening skills

Several researchers have shown that even skilled listeners do not always perform in superior ways. For example, in Cook's (1987) ground-breaking experiment, even music students could not reliably discern between the correct version of a piece with its recapitulation in the fundamental key and a manipulated version with the recapitulation in a distant key. Kopiez and Platz (2009) found that only half of their subjects (music students) noticed a clash of keys between a song's melody and accompaniment in a non-attentive listening situation. Attention and conscious processing of certain aspects are probably necessary to elicit expert listening performance.

Literature exists on the evaluation of performance (e.g. Ziv & Moran, 2006), improvisations (e.g. Eisenberg & Thompson, 2003),

and compositions (e.g. Leung, Wang, Lee & 2009), but to our knowledge, no comparisons have been drawn between improvised and composed pieces. We will therefore later discuss if there are perceptually salient cues that can differentiate generative types.

In an informal pilot study to this paper we presented two pieces of music to an audience of experts who had attended a talk at an ESCOM conference (see Lehmann & Kopiez, 2002). Listeners were asked to judge whether the examples were improvised or composed. Surprisingly, the votes were roughly equally distributed between the response alternatives. The present study replicates our 2002 pilot study under laboratory conditions and attempts to investigate if the generative types of composed and non-composed music can be identified by expert listeners when they are not consciously searching for them.

2. METHOD

2.1 Participants and procedure

Participants. Participants in this listening experiment were 102 college music students of typical undergraduate age and experience. They were enrolled in music psychology or systematic musicology courses and were tested in groups.

Procedure. Listeners were presented six musical excerpts in randomized order. The listeners completed a short questionnaire after hearing each excerpt. On 5-point unipolar rating scales they marked their agreement to adjectives “coherent/orderly”, “spontaneous/improvised”, and “deliberated/rehearsed”. They also indicated whether they knew the particular piece and liked it. Retest-reliability was assessed with a subsample of participants.

2.2 Musical examples

The stimuli consisted of three pairs of two pieces. Each pair was made up of a carefully selected composed piece and its matched non-composed or improvised mate (see Table 1; supplementary audio material). Excerpts were cut to a length of about one minute.

Style and composer	Generative type	
	Composed	Non-composed (“Improvised”)
Classical (Mozart)	Minuet KV 282 (Brendel, 1992)	Dice game KV294d/516f; Op. posthumus (Reuter, 2001)
Romantic (Scriabin)	Sonata No. 9, Op. 68 (Ashkenazy, 1989)	Improvisation on Scriabin motives (Philipp, 1999)
Free Jazz (Rissi et al., 1996)	“Papago” [composed section]	“Papago” [improvised section]

Table 1: Musical examples used in the listening experiment (excerpts of approx. one minute).

Note: Some of the examples are available as additional material to this short paper.

3. RESULTS

3.1 Interrelation of dependent variables

Correlations between the three dependent variables (the scales coherent/orderly, spontaneous/improvised, and deliberated/rehearsed) were all significant and of expected direction. Spontaneous/improvised was negatively correlated with coherent/orderly ($r = -.59$) and deliberated/rehearsed ($r = -.68$). Coherent/orderly and deliberated/rehearsed were positively associated ($r = .60$). We averaged coherent/orderly and deliberated/rehearsed to compute a more stable measure which we called “coherent/rehearsed”. The resultant variable was highly negatively correlated with spontaneous/improvised ($r = -.71$). Our dependent variables in the following analyses were *spontaneous/improvised* and *coherent/rehearsed*, which appear to capture opposing characteristics of our stimuli.

A subsample of 40 listeners received repeated exposure to one of the pieces (Scriabin, composed). The retest-reliability was moderate with correlations of .40 to .52. The reliabilities were reliable, not high, but still acceptable.

3.2 Listeners detected difference in free jazz but not in classical and romantic examples

The next analyses assessed the influence of the within-subject variables *generative type* (levels: composed, non-composed) and *style* (levels: Romantic, Free Jazz), separately for our two dependent variables (coherent/rehearsed and spontaneous/improvised). The Mozart pieces had to be excluded due to lack of variability on some scales. The generative type was of specific interest because an effect here would show that listeners detected differences in structure or performance. The directed hypothesis was that listeners would experience improvisations as more spontaneous/improvised while compositions would appear more coherent/rehearsed.

The mean *coherent/rehearsed* scores for the Scriabin and Rissi pieces for the “generative type” factor differed significantly beyond the .01 level: $F(1,101) = 86.3, p < .001$, with a large effect size of partial eta-squared = .46 (see Tables 3 and 4). The mean ratings for the effect of the “piece” factor also differed significantly at the .01 level: $F(1,101) = 315.5, p < .001$, yielding a large effect with partial eta-squared of .76. The “piece by generative type” interaction was also reliable, $F(1,101) = 40.2, p < .001$, showing a partial eta-squared of .28 (see Figure 1).

The mean *spontaneous/improvised* scores for the Scriabin and Rissi pieces for the “generative type” factor differed significantly beyond the .01 level: $F(1,101) = 21.5, p < .001$, with a large effect size of partial eta-squared = .17 (see Tables 5 and 6). Also, the mean ratings for the “style” effect differed significantly at the .01 level: $F(1,101) = 142.3, p < .001$, yielding a large effect with partial eta-squared of .58. And the “piece by generative type” interaction

was reliable, $F(1,101) = 21.1$, $p < .001$, showing a partial eta-squared of .17 (see Figure 2, for more details Lehmann & Kopiez, in press).

Using a paired sample t-test, differences between composed and non-composed versions for the spontaneous/improvised and the coherent/rehearsed scales were analyzed. As Figures 1 and 2 illustrate, the results for Scriabin were not reliable. However, the results were significant for Rissi ($t[101] = 5.9$, $p < .01$ for spontaneous/improvised and $t[101] = 10.8$, $p < .01$ for coherent/rehearsed). Thus, the significant “generative type” and “generative type * style” effects were solely due to the differences for the Rissi piece. Finally, knowledge of the piece as a covariate did not lead to significant results, suggesting that knowing the piece did not matter.

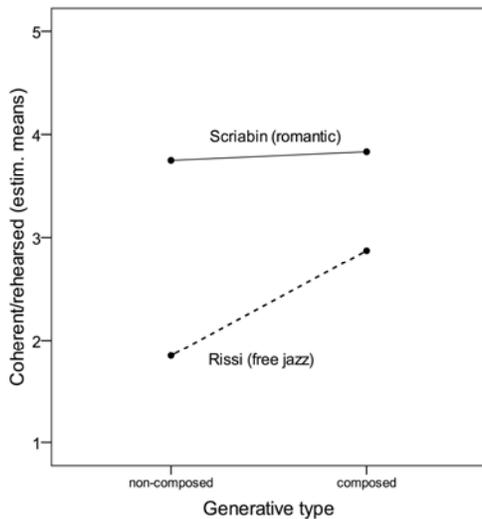


Figure 1: Reliable interaction effect between style and generative type for the dependent variable *coherent/rehearsed*.

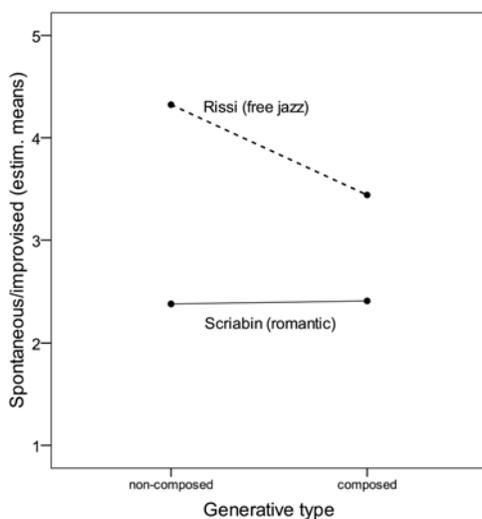


Figure 2: Reliable interaction effect between style and generative type for the dependent variable *spontaneous/improvised*.

For the Mozart piece we analyzed the effect of generative type using non-parametric statistics with two separate Mann-Whitney tests, one for each dependent variable. The mean ratings for *coherent/rehearsed* for the non-composed version ($M = 4.70$, $SD = .47$, $95\% \text{ CI} = 4.61/4.80$) were not reliably higher than those for the composed version ($M = 4.69$, $SD = .61$, $95\% \text{ CI} = 4.56/4.80$). Similarly, non-significant results were obtained for the scale *spontaneous/improvised* for the non-composed ($M = 1.20$, $SD = .53$, $95\% \text{ CI} = 1.09/1.30$) and for the composed version ($M = 1.35$, $SD = .83$, $95\% \text{ CI} = 1.18/1.51$). Thus, the results of the Mozart comparison follow closely those found for the Scriabin pieces and would lead to a similar graphical representation (see Scriabin in Figure 1 and 2 for illustration).

In sum, our main dependent variables, spontaneous/improvised and coherent/rehearsed, were highly negatively correlated. Our listeners distinguished readily between musical styles, but did not assign differential ratings to matched composed and non-composed versions of the classical (Mozart) and romantic (Scriabin) music pieces. The free jazz (Rissi) examples, both from exactly the same piece, were rated differentially.

4. DISCUSSION

Our assumption was that the two generative types (or processes) leading to composition and improvisation are much more akin than commonly assumed and lead to products that are perceptually difficult to keep apart. Obviously, trying to show that expert listeners cannot easily discern one generative process from the other is epistemologically questionable because we were essentially aiming for a null result.

That the listeners would be sensitive to the different styles was expected and confirmed. More interesting was the fact that the distinction of generative types – that is, composed versus non-composed – for our classical (Mozart) and romantic (Scriabin) examples was not significant. Identification appears difficult, if not impossible, when the non-composed stimulus is of high quality and played on a single instrument. Yet, listeners succeeded in the case of the free jazz excerpts (Rissi): The improvised version was judged to be reliably more spontaneous/improvised, the composed one more coherent/rehearsed. What could be the reason for this perceptual reality?

Based on our exploratory study, we propose the following explanation: Listeners do not have clear categories for improvisation and composition, but they can search for cues (Deliège, 2006) that may hint at the underlying generative types or processes. However, contrary to what musicologists or theorists might think in the aesthetic tradition of Western music, the cues do not reside in the structural features of the work itself but rather emerge from its performative embodiment. Improvisation implies a freedom which is particularly instantiated in the rhythmic execution of the music. More “togetherness” and precision of an ensemble may indicate composition, while a higher degree of entropy could signal improvisation. This cue is more easily extracted from performances with several instruments as was the case in the free jazz example: When played from notation, the ensemble will sound more together, whereas during the improvisation the rhythms are more freely executed.

Future research is necessary. Our tentative explanation requires further experimentation in which rehearsed pieces are presented with varying degrees of freedom in performance. Given the limited scope of our choice of pieces and their duration, a next step could also be the production of further and longer stimuli by asking performers to play a piece and then improvise on it. For example, performers of early music have internalized idiomatic features of a given style and can skillfully generate works in real-time that, to the uninitiated listener, are virtually indistinguishable from compositions. Finally, it may be useful to include audio-visual material to broaden the scope to include ensemble performance, because improvisation in groups requires (facial and bodily) communication.

5. ADDITIONAL FILES

Rissi_IMP.mp3 - Musical example of Free Jazz (improvised section)

Rissi_COM.mp3 - Musical example Free Jazz (composed section)

Scriabin_IMP.mp3 - Musical example Romantic style (improvised section)

Scriabin_COM.mp3 - Musical example Romantic style (composed section)

6. REFERENCES

- Ashkenazy, V. (1989). *Scriabin piano sonatas* [CD]. London: Decca.
- Auslander, P. (in press). Jazz improvisation as a social arrangement. In N. Cook & R. Pettengill (eds.), *Music as performance: New perspectives across the disciplines*. Ann Arbor: Michigan University Press.
- Brendel, A. (1992). *Mozart: Sonatas KV533/494 & 282* [CD]. Hollywood: Philips Digital Classics.
- Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception*, 5(2), 197-206.
- Deliège, I. (2006). Creative support to elaborate a model of music listening. In I. Deliège & G. A. Wiggins (eds.), *Musical creativity: Multidisciplinary research in theory and practice* (pp. 63-77). Hove: Psychology Press.
- Eisenberg, J., & Thompson, W. F. (2003). A matter of taste: evaluating improvised music. *Creativity Research Journal*, 15(2/3), 287-296.
- Fox, C., & Osmond-Smith, D. (2009). "Scelsi, Giacinto." In *Grove Music Online*. Oxford Music Online, (accessed December 27, 2009).
- Kopiez, R. & Platz, F. (2009). The role of listening expertise, attention, and musical style in the perception of clash of keys. *Music Perception*, 26(4), 321-334.
- Lehmann, A. C. & Ericsson, K. A. (1998). Historical developments of expert performance: Public performance of music. In A. Steptoe (ed.), *Genius and the mind* (pp. 67-94). Oxford: Oxford University Press.
- Lehmann, A. C. & Kopiez, R. (2002). Revisiting composition and improvisation with a historical perspective. In M. Britta, & M. Mélen (eds.), *Proceedings of the 10th Anniversary Conference of the European Society for the Cognitive Sciences of Music (ESCOM): Musical Creativity*, Liège, Belgium, 5-8 April [CD-ROM].
- Lehmann, A. C. & Kopiez, R. (in press). *The difficulty of discerning between composed and improvised music*. Musicae Scientiae.
- Lehmann, A. C. (2005). Komposition und Improvisation: Generative musikalische Performanz [Composition and improvisation: Generative musical performance]. In T. H. Stoffer & R. Oerter (eds.), *Allgemeine Musikpsychologie* (pp. 913-954). Göttingen, Germany: Hogrefe.
- Leung, C., Wang, Y., Lee, A. (2009). Assessment of undergraduate students' music compositions. *International Journal of Music Education*, 27, 250-268.
- Nettl, B. (1998). Introduction: An art neglected in scholarship. In B. Nettl & M. Russell (eds.), *In the course of performance* (pp. 1-23). Chicago: Chicago University Press.
- Philipp, G. & Pruggmayer-Philipp, U. (1999). *Klavierimprovisationen* [CD]. Berlin: Eterna.
- Pressing, J. (1998). Psychological constraints on improvisational expertise and communication. In B. Nettl & M. Russell (eds.), *In the course of performance* (pp. 47-68). Chicago: Chicago University Press.
- Reuter, C. (2001). *Musikalische Würfelspiele von Mozart, Haydn und anderen großen Komponisten* [Musical dice games by Mozart, Haydn and other great composers] [CD-ROM]. Mainz: Schott.
- Rissi, M., Mazzalo, G., & Geisser, H. (1996). *Fuego* [CD]. Root, Switzerland: Creative Works.
- Sidran, B. (1995). *Talking jazz. An oral history*. New York: Da Capo.
- Ziv, N., & Moran, O. (2006). Human versus computer: The effect of a statement concerning a musical performance's source on the evaluation of its quality and expressivity. *Empirical Studies of the Arts*, 24(2), 177-191.