

Modelling the relationships between emotional responses to, and musical content of, music therapy improvisations

Psychology of Music

Psychology of Music

Copyright © 2008

Society for Education, Music

and Psychology Research

vol 36(1): 25–45 [0305-7356

(200801) 36:1; 25–45]

10.1177/0305735607079714

<http://pom.sagepub.com>

GEOFF LUCK, PETRI TOIVIAINEN, JAAKKO ERKKILÄ,
OLIVIER LARTILLOT AND KARI RIIKKILÄ

UNIVERSITY OF JYVÄSKYLÄ, FINLAND

ARTO MÄKELÄ

SATAKUNTA DISTRICT OF SERVICES FOR THE INTELLECTUALLY DISABLED,
FINLAND

KIMMO PYHÄLUOTO

PÄÄJÄRVI FEDERATION OF MUNICIPALITIES, FINLAND

HEIKKI RAINE

RINNEKOTI-FOUNDATION, FINLAND

LEILA VARKILA

PÄÄJÄRVI FEDERATION OF MUNICIPALITIES, FINLAND

JUKKA VÄRRI

SUOJARINNE FEDERATION OF MUNICIPALITIES, FINLAND

ABSTRACT This article reports a study in which listeners were asked to provide continuous ratings of perceived emotional content of clinical music therapy improvisations. Participants were presented with 20 short excerpts of music therapy improvisations, and had to rate perceived *activity*, *pleasantness* and *strength* using a computer-based slider interface. A total of nine musical features relating to various aspects of the music (timing, register, dynamics, tonality, pulse clarity and sensory dissonance) were extracted from the excerpts, and relationships between these features and participants' emotion ratings were investigated. The data were analysed in three stages. First, inter-dimension correlations revealed that ratings of *activity* and *pleasantness* were moderately negatively correlated, *activity* and *strength* were strongly positively correlated, and *strength* and *pleasantness* were moderately negatively correlated. Second, a series of cross-correlation analyses revealed that the temporal lag between musical features and listeners' dimension ratings differed across both variables and dimensions. Finally, a series of linear regression analyses produced significant feature prediction models for each of the three dimensions, accounting for 80 percent (*activity*), 57 percent (*pleasantness*) and 84 percent (*strength*) of the variance in participants' ratings. *Activity* was best predicted by high note density and high pulse clarity, *pleasantness* by low note density and high tonal clarity, and *strength* by high mean velocity and low note density. The results are discussed in terms of their fit with

other work reported in the music psychology literature, and their relevance to clinical music therapy research and practice.

KEYWORDS: *continuous response, perceived emotion*

Introduction

In the early 1960s, Paul Nordoff and Clive Robbins, two therapists working with handicapped children, began to develop a form of music therapy in which improvisation played a central role (Nordoff and Robbins, 1965). The use of improvisation in music therapy has been developed over the years, and, since Nordoff and Robbins's pioneering work, improvisation has been used in music therapy with a broad range of clients suffering from a variety of clinical conditions (see Bruscia, 1987 for a review of improvisational models of music therapy and Wigram, 2004, for a more recent discussion of improvisational music therapy methods and techniques).

Improvisation can be used as the sole method in music therapy, or it can be connected to playing or listening to pre-composed music, other musical activity, or other non-musical activity, such as visual art or discussion. Whichever approach is taken, there is always musical and non-verbal communication present in the improvisational situation through which the therapist tries to communicate with the client, and encourage them to act in a desired way. Goals of an improvisation session are typically related to the client's physiological, cognitive, emotional or social functioning.

Traditionally, music therapy improvisation has been based around the use of acoustic instruments and the human voice. More recently, however, the use of electric and electronic instruments has become more common. Electronic keyboards, for example, may be used for economic or youth culture-related reasons. Moreover, the implementation of MIDI instruments allows the computational analysis of improvisations with relative ease in comparison to the analogue data produced by acoustic instruments.

As mentioned above, one of the key areas in which improvisational music therapy is used is related to changes in a client's emotional functioning. Thus, the connection between musical expression and perceived emotional meaning is essential. Indeed, most improvisational models of music therapy emphasize the role of improvisation in exploring and expressing emotions (see Bruscia, 2001). However, detailed investigation of the relationship between emotional (personal) experience and musical event(s) in improvisation is considered a very challenging task in the music therapy literature (see Smeijsters, 1997). This is true particularly with regard to free improvisation, where the client creates sounds and music without any direction, guidelines or rules provided by the therapist (definition by Bruscia, 1987). This type of music is seldom successfully described with terms originating in traditional music analysis. It is perhaps for this reason that Vink (2001) has suggested that musical material is more relevant for music psychologists (who tend to study composed music) than music therapists.

Nevertheless, music therapists are becoming increasingly interested in studying the connections between listeners' emotional experience and the musical material of improvisation (e.g. Lee, 2000; De Backer, 2004; Erkkilä, 2004). Moreover, the need for multidisciplinary collaboration has been stressed. For example, Wosch (2003)

points out that in order to improve improvisation analyses, music therapists should make more use of the knowledge and expertise of music psychologists and emotion psychologists. Due to the often minimal presence of predetermined musical referents in free improvisation (see Pavlicevic, 2000), there is a need to be able to capture the most essential musical features – whatever they are – and connect them to the psychological meanings, especially those relating to emotional content, that they represent. In other words, there is a need to be able to define and extract the clinically relevant combinations of musical features that are ‘hiding’ within the improvisations.

However, these combinations of features change and evolve as an improvisation unfolds, raising the question of how best to investigate their relationship to a listener’s perception of emotional content. One solution is to take continuous ratings of listener’s perceptions of emotional content and examine the relationship between these ratings and the changing combinations of musical features present in the improvisation. This methodology is used widely in studies of emotion and music because of its dynamic nature of data-collection – participants’ responses can be collected throughout the duration of a temporally extended stimulus, such as a musical passage.

The continuous response methodology involves participants giving real-time responses to stimuli using some sort of physical or computer-based slider device. It is most frequently used for eliciting responses along a single scale, although there are examples of the concurrent use of two scales (e.g. Schubert, 2004). When a single scale is used, the two extremes of the scale are labelled with the two extremes of the dimension under investigation and participants are required to move a pointer between these two extremes to indicate their real-time response to the stimuli with which they are presented.

Examples of the use of this method in the literature include the investigation of judgments of tension (Madsen and Fredrickson, 1993; Fredrickson, 1995, 1999; Krumhansl, 1996, 1997; Krumhansl and Schenck, 1997; Fredrickson and Coggiola, 2003; Toiviainen and Krumhansl, 2003), memorability and openness (Krumhansl, 1998), the amount and quality of emotions (Krumhansl, 1997, 1998; Krumhansl and Schenck, 1997; Schubert, 1999, 2004), aesthetic response (Coggiola, 2004), mood state (Goins, 1998), arousal (Madsen, 1997; Schubert, 2004), and affect (Madsen, 1997). The most relevant of these to the present study are the studies which have examined emotion-related judgments.

In a study of listener’s perception of musical emotions, Krumhansl (1997) presented participants with six excerpts of classical music, two each to represent *sadness* (Tomaso Albinoni: Adagio in G minor for Strings and Orchestra; Samuel Barber: Adagio for Strings, Op. 11), *fear* (Gustav Holst: ‘Mars – the Bringer of War’ from *The Planets*; Modest Mussorgsky: *Night on the Bare Mountain*), and *happiness* (Antonio Vivaldi: ‘La Primavera’ (Spring) from *The Four Seasons*; Hugo Alfven: *Midsommarvaka*). Participants listened to the excerpts and rated their perceptions of these three emotions, one at a time, using a computer mouse to control an on-screen slider. At the end of each excerpt, participants completed a series of questionnaire items in which they rated how they felt while listening in terms of 13 different emotions, the pleasantness and intensity of the music, and their familiarity with the piece.

Both types of data were subjected to factor analysis, and the resulting two-factor solutions indicated similar patterns of factor loadings for both the continuous emotion ratings and the questionnaire-based emotion ratings. In other words, listener’s dynamic

ratings of emotion perceived in the excerpts matched the expected emotional representation of each excerpt. As regards the continuous data, the ratings of emotion remained at a fairly high level throughout each excerpt, but exhibited local variation. This suggests that although the overall feel of a piece of music may induce the perception of a particular emotion quite strongly, structural features of the music may increase or decrease the perception of emotion as the music unfolds. Indeed, Schubert (2004) asserts that, since in studies of music and emotion there tends to be a high level of agreement between listeners as to the emotion expressed by a particular piece of music, a significant proportion of emotional expression in music must be related to the musical features of a particular piece. Nonetheless, Krumhansl (1997) made no attempt to examine the relationship between musical features of the excerpts and participants' ratings in her study.

Krumhansl and Schenck (1997), however, investigated the structural and expressive mappings between music and dance, specifically between Mozart's Divertimento No. 15 and George Balanchine's choreography specifically written to accompany the piece. Participants were presented either with an auditory stimulus (an audio-recording of the piece), a visual stimulus (a video-recording of a dancer performing Balanchine's choreography) or an audio/visual stimulus (both the music and the dance performance). One of the tasks required of participants was to rate the amount of emotion expressed in the stimuli they were presented with. To do this, participants depressed a controller pedal attached to a digital keyboard, and the position of the controller was sampled at 250 ms intervals.

Krumhansl and Schenck (1997) found that perceived emotion ratings tended to increase after new material was introduced and decrease towards section ends. Furthermore, ratings tended to follow this pattern regardless of the stimulus modality. Thus, ratings of emotion were found to correlate with the structural features of the music or dance performance.

In another study of the relationships between musical features and listeners' perception of emotion, Krumhansl (1998) presented participants with audio recordings of the first movements of two chamber music works for strings (Mozart's String Quintet in C Major, and Beethoven's String Quartet in A minor), and in both cases participants had to rate the amount of emotion they perceived as the pieces unfolded. As in Krumhansl's (1997) study, participants indicated their responses by using a computer mouse to control an on-screen slider, the position of which was recorded at 250 ms intervals. In line with Krumhansl and Schenck's (1997) findings, the amount of perceived emotion was found to be related to structural features of the two pieces.

However, one shortcoming of these studies is that they only measured perceptions of emotion on a rather general level, making no distinction between the different types of emotions which might be expressed, nor relating their findings to different theories regarding the structure of emotions. One such group of theories, known as dimensional theories of emotion, have been suggested to be particularly well-suited to studies that examine the dynamic changes in emotional expression during a piece of music (Juslin and Sloboda, 2001).

Dimensional theories of emotion hold that emotional meaning can be described within a multidimensional emotion space comprised of a small number of dimensions, most frequently cited as relating to *valence/pleasantness*, *activity/arousal* and *potency/strength* (e.g. Osgood et al., 1957). Each dimension is assumed to be anchored by

semantic terms representing polar opposites, such as happy–sad for *valence*, active–inactive for *activity* and weak–strong for *potency*. Schubert (2001) notes that references to the first two of these dimensions are frequently found in the music-emotion literature. *Potency*, however, is somewhat less frequently described, and its role in the emotion space not so clearly defined.

Schubert (2004) examined the relationship between musical/psychoacoustic features and different emotional dimensions, using a continuous response methodology, and time-series analyses. Specifically, he investigated the relationship between five musical features (melodic contour, tempo, loudness, texture and timbral sharpness) and participants' concurrent ratings of perceived valence and arousal. Participants used a computer mouse to indicate their ratings in a two-dimensional emotion space (2DES) in which the x axis represented valence (happy–sad), and the y axis represented arousal (aroused–sleepy). Schubert recorded participants' responses at 1 s intervals, and found that ratings of arousal were positively related to loudness, and, to a lesser extent, tempo. Meanwhile, valence was found to be somewhat related to melodic contour, although the results for this dimension were not conclusive.

Thus, Schubert's study showed that it is possible to extract dynamic musical features using computational methods and examine how they relate to listeners' ratings of at least two dimensions of emotion, namely perceptions of valence and arousal. However, Schubert's study is not without its problems.

First, although Schubert argued that it was possible for listeners to rate two emotional dimensions concurrently, the present authors feel that this requirement may well have overwhelmed participants. With participants having to divide their attentional resources between listening to the music, rating their perceptions of valence and rating their perceptions of arousal, it seems unlikely that they would be able to provide data that accurately reflected their perceptions.

Second, there is the issue of the time lag between the musical features and participants' responses to them. Schubert calculated that this tended to vary between 1 s and 3 s, with sudden changes in loudness giving rise to the shortest lag. Three seconds seems like a significant lag between the occurrence of a musical feature and participants' response to it. Might this not indicate that, as suggested above, that the task required of participants was too demanding?

A more general limitation of previous studies that have examined the relationships between listeners' perceptions of emotion in music using the continuous response method is that they have all used composed music of the western classical tradition as stimuli. None of them have examined other forms of music-making, such as improvisation. Thus, it is not clear whether the connections between composed music and perceptions of emotion highlighted by previous studies also hold for improvised music.

One advantage of using improvised music as a stimulus is that the issue of learnt associations is minimized. Music often becomes associated with memories of events or contexts, such as when you first heard a particular piece of music, or who you were with at the time, and the music then serves as a trigger to recall that event or context. A well-known example of this type of association is the so-called 'Darling, they're playing our tune' theory (Davies, 1978).

Such associations could be minimized still further by using genre-free improvisations, such as those produced in music therapy sessions, as opposed to those played in

a particular style, e.g. jazz, classical. Improvisations produced by a typical individual with mental retardation are not well structured, or of 'high quality', in the accepted sense of the term. Such improvisations do not represent western classical or any other public performance music tradition.

The use of improvised music free of these kinds of associations, therefore, would allow one to concentrate primarily on the effects of the musical features upon listeners' perceptions of emotion. Of course, any kind of musical stimulus may invoke a learnt association in a listener. For example, the general 'style' of the piece, or a particular chord progression (if indeed chords were played), may have associations. Such associations, however, are hard to avoid, but the use of improvised music of this type at least helps keep them to a minimum.

The aim of the present study was to develop previous work in three different ways. First, to apply the continuous response methodology to the investigation of improvised music. Second, to develop the way this methodology is applied, taking separate ratings of the three fundamental dimensions of emotion, in the present study labelled *activity*, *pleasantness* and *strength*. Third, to examine the relationships between musical features and perceptions of emotion in more detail by extracting a much larger number of dynamic musical features from the music compared to previous studies.

Participants were presented with clinical music therapy improvisations, and asked to provide continuous ratings of perceived *activity*, *pleasantness* or *strength*. Note that participants were asked to rate the emotion that they felt the music was trying to express, rather than the emotional response they might feel. This is because it has been suggested that it is easier to agree on the emotion expressed by music than the emotion evoked in listeners (e.g. Campbell, 1942; Hampton, 1945; Swanwick, 1973). The relationship between participants' rating and computationally extracted musical features was investigated using linear regression analyses.

A total of nine musical features (note density, articulation, mean pitch, standard deviation (SD) of pitch, mean velocity, SD of velocity, tonal clarity, pulse clarity and dissonance) were extracted from the improvisation excerpts, and their relationship to ratings of perceived *activity*, *pleasantness* and *strength* investigated using linear regression. Predictions regarding specific relationships were hard to formulate because most of the extracted features had not been investigated in this way before. Another problem was the lack of consistent emotion-related terms used in previous work. Nonetheless, the following predictions were tentatively made, based largely upon Gabrielsson and Lindström's (2001) detailed review of the literature relating to the influence of musical structures on musical expression (see the notes to this article for terminology clarification. Note that we do not suppose a 1:1 relationship between our terminology and that used by the cited authors, only that our terms are somewhat analogous to theirs).

It was predicted that ratings of *activity* would be positively related to more detached articulation (Wedin, 1972), higher mean pitch (Scherer and Oshinsky, 1977), larger SD of pitch¹ (Scherer and Oshinsky, 1977), higher mean velocity² (Schubert, 2004), smaller SD of velocity³ (Scherer and Oshinsky, 1977), and higher levels of dissonance (Costa et al., 2000). It was predicted that ratings of *pleasantness* would be positively related to lower mean pitch (Scherer and Oshinsky, 1977), larger SD of pitch (Scherer and Oshinsky, 1977), smaller SD of velocity (Scherer and Oshinsky, 1977), and lower levels of dissonance (Wedin, 1972; Costa et al., 2000). Finally, it was predicted that ratings of *strength* would be positively related to higher mean pitch (Scherer and Oshinsky, 1977;

Costa et al., 2000), higher mean velocity⁴ (Kleinen, 1968) and higher levels of dissonance (Costa et al., 2000).

Method

LISTENING TEST

Participants

Twenty-five undergraduate students from the University of Jyväskylä took part in the experiment, and were awarded course credit for their participation. All participants were enrolled on one of the music department's three undergraduate programmes (musicology, music education or music therapy), and were thus deemed as being musically experienced.⁵

Stimuli

Stimuli used in the experiment were 20 randomly selected 1-minute excerpts of the client's part of a corpus of full-length client–therapist improvisations collected by professional music therapists throughout Finland.⁶ The therapists used a set of two identical 88 key weighted action MIDI keyboards (Fatar Studiologic SL-880 PRO Master keyboard) to improvise with their clients in their regular music therapy sessions (the therapist improvised on one keyboard while the client improvised on the other). The clients were individuals with a developmental disability or a psychiatric diagnosis, and consent was obtained to allow their improvisations to be used anonymously for research purposes. All improvisations were performed with similar volume settings, using the same MIDI grand piano voice, and were recorded by the therapist using the Cubase sequencer software (manufacturer: Steinberg). The therapist's and client's parts were recorded on separate MIDI tracks. Only the clients' improvisations were used in the present study.

Continuous response recording procedure

Participants were asked to provide continuous ratings of three identical blocks of 20 excerpts, on each occasion rating their perceptions of *activity*, *pleasantness* or *strength*. A 'slider' interface was developed in the Max/MSP graphical environment (manufacturer: Cycling '74), which presented the 20 excerpts in a single block, and recorded participants' continuous responses at 500 ms intervals, on a PC running Microsoft® Windows® XP (the interface was in fact designed as a Max/MSP external object that could be run either on an Apple Macintosh® or PC). The block of excerpts was presented three times to each participant, and the order in which participants were required to rate the three dimensions was counterbalanced. The slider interface used the PC's internal MIDI synthesizer to present the 20 musical excerpts using the default piano sound.

A large horizontal slider, moved with the mouse, occupied the main part of the interface window. Three versions of the interface were created, corresponding to the three dimensions to be rated: *activity*, *pleasantness* and *strength*. Each version indicated the actual musical meaning associated with the extremities of the slider: respectively, from the right to the left of each slider, 'active' (*aktiivinen* in Finnish)/'inactive' (*ei aktiivinen*), 'pleasant' (*miellyttävä*)/'unpleasant' (*epämiellyttävä*) and 'strong' (*voimakas*)/'weak' (*heikko*). The middle of each slider represented the neutral value between the two extremes.

Participants were also provided with information regarding the progression of the experiment: above the slider they were shown the current excerpt number they were listening to (out of 20), and to the right of the slider they were shown the temporal progression of the excerpt they were currently listening to. Figure 1 shows the interface window corresponding to the activity measurement.

Since the rating of each successive excerpt may be influenced by the previous excerpt (participants might, for example, rate the beginning of each excerpt relative to the ending of the previous excerpt), the 20 excerpts were presented in a random permutation, computed automatically every time a new block of stimuli was presented. The rating of the 20 excerpts was stored in the original non-permuted order in each experiment record. The interface program also stored the permutation order in case this parameter was of interest. This parameter was not, however, taken into consideration in the procedure presented in this article.

The first excerpt was preceded by a 10 s pause, indicated by an on-screen countdown, and all subsequent excerpts were separated by a 10 s pause/countdown, during which the slider was reset back to the middle. Each excerpt was 60 s long, and the position of the slider was recorded at 500 ms intervals, resulting in a 120 sample-long record for each excerpt, for each participant. At the end of each block, a file was produced that contained the 20 values of the permutation order, followed by the 20 continuous measurements. Figure 2 shows the activity measurement of one participant. Before the

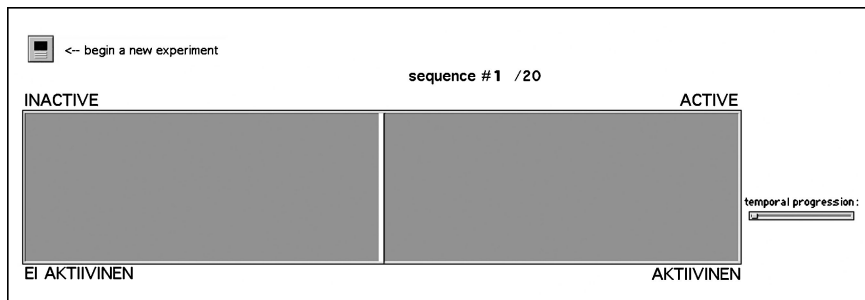


FIGURE 1 *The slider interface for activity. The extremities of the sliders for the other two dimensions (pleasantness and strength) were anchored by 'pleasant' (miellyttävä) and 'unpleasant' (epämiellyttävä), and 'strong' (voimakas) and 'weak' (heikko), respectively. For all three dimensions, the current excerpt number (out of 20) was displayed above the slider, and the temporal progression of the current excerpt was displayed to the right of the slider.*

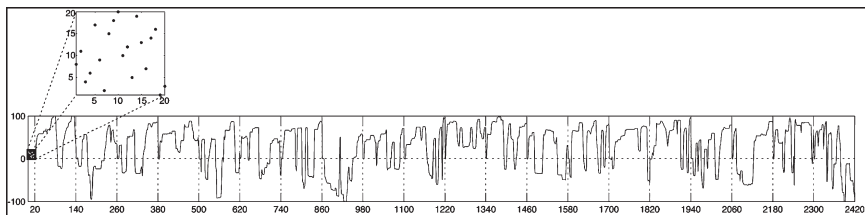


FIGURE 2 *Perceived activity ratings of a single participant, for all 20 excerpts. The beginning of each excerpt is delineated by a vertical dotted line. Note that the first 20 data points are taken up by the zoomed area, which shows the ordering of the 20 excerpts for this particular participant.*

actual experiment, participants were presented with three practice trials using non-experimental stimuli to familiarize them with the experimental task and interface. Participants' responses to these practice trials were not recorded.

MUSICAL FEATURE EXTRACTION

The musical stimuli used in the listening experiment were subjected to a computational analysis to obtain a set of quantitative descriptors representing a variety of musical features. The analysis was carried out from the MIDI file representation with algorithms implemented for the purpose of this study in MATLAB using the MIDI Toolbox (Eerola and Toiviainen, 2004). To allow comparison with the continuous rating data, the analysis was carried out using a sliding window of fixed length (6 s). For each window, the temporal point to which the values of the musical variables were associated was the end point of the window. The temporal location of the window's end point was changed from 0.5 s to 60 s with steps of 0.5 s, measured from the first note onset of each stimulus. This resulted in a time series of 120 points for each variable and each stimulus.

The choice of a 6 s window was based upon the fact that estimation of auditory sensory memory varies from 3 to 8 s (e.g. Treisman, 1964; Darwin et al., 1972; Fraisse, 1982). Preliminary analyses, trying various window lengths, indicated that shorter lengths resulted in discontinuities in the data, and increased jitteriness, while longer lengths smoothed the data too much. A 6 s window was thus a compromise between these two extremes.

The musical features to be extracted were chosen on the basis of the following criteria. First, the features had to be extractable from the information available in the MIDI file format, i.e. from note onset and offset times, pitches in semitones and key velocity. Second, they had to comprise several musical dimensions in order to provide a comprehensive representation of the musical content. Finally, they had to encompass features with differing levels of complexity, ranging from psychophysical features, such as note density and dynamics, to more context-dependent features, such as pulse clarity and tonality. In what follows, each of the musical feature variables used in the analysis is described.

A. Temporal surface features

1. *Note density*: Number of notes divided by the length of the window.
2. *Articulation*: Proportion of temporal intervals during which at least one note is being played. Values close to unity indicate legato playing, while values close to zero indicate staccato or a substantial proportion of silent periods.

B. Features related to register

These features were based on the MIDI pitch values of notes.

3. *Average pitch*.
4. *SD of pitch*.

C. Features related to dynamics

These features were based on the note-on velocity values.

5. *Average note-on velocity*.
6. *SD of note-on velocity*.

D. Features related to tonality

These features were based on the Krumhansl-Schmuckler key-finding algorithm (Krumhansl, 1990).

7. *Tonal clarity*: To calculate the value of this feature, the pitch-class distribution of the windowed stimulus was correlated with the 24 key profiles representing each key (12 major keys and 12 minor keys). The maximal correlation value was taken to represent tonal clarity.

E. Other features

8. *Pulse clarity*: To calculate the value of this variable, a temporal function was first constructed by summing Gaussian kernels located at the onset points of each note. The height of each Gaussian kernel was proportional to the duration of the respective note; the SD was set to 50ms (see Toiviainen and Snyder, 2003). Subsequently, the obtained function was subjected to autocorrelation using temporal lags between 250 ms and 1500 ms, corresponding to commonly presented estimates for the lower and upper bounds of perceived pulse sensation (Warren, 1993; Westergaard, 1975). To model the dependence of perceived pulse salience on beat period, the values of the autocorrelation function were weighted with a resonance curve having the maximal value at the period of 500ms (Toiviainen 2001; see also Van Noorden and Moelants, 1999). The maximal value of the obtained weighted autocorrelation function was taken to represent the degree of instantaneous pulse clarity.
9. *Sensory dissonance*: Musical dissonance is partly founded on cultural knowledge and normative expectations, more suitable for the analysis of improvisation by expert rather than by non-expert musicians. More universal is the concept of *sensory dissonance* (Helmholtz, 1877), which is related to the presence of beating phenomena caused by frequency proximity of harmonic components. Sensory dissonance caused by a couple of sinusoids can be predicted simply. The global sensory dissonance generated by a cluster of harmonic sounds is then computed by adding the elementary dissonances between all the possible couples of harmonics (Plomp and Levelt, 1965; Kameoka and Kuriyagawa, 1969). In the present study, the same instrumental sound (MIDI default piano sound) was used during the improvisations and the listening tests, and, therefore, also in the dissonance measure. Following a detailed spectral analysis, the spectral component of this piano sound was modelled by selecting the six first harmonics of each note and assigning their successive amplitudes to a geometric progression of common ratio 0.8. The amplitude envelope decay of each note was modelled as a negative exponential, with a time constant linearly related to pitch height.

Results

INTER-SUBJECT AND INTER-DIMENSION CORRELATIONS

Before analysing the data, inter-subject correlations were calculated to see whether it would be justifiable to use the mean rating of each of the three dimensions in subsequent analyses. A preliminary visual inspection of the data revealed that one participant's ratings were markedly different from all others, and it was subsequently discovered that this participant's background and experience were not in keeping with the relative homogeneity of the other 24 participants. Specifically, this participant was

familiar with music therapy improvisation material, and experienced in music therapy analysis. This participants' data was thus excluded from the calculation of inter-subject correlations, and any further analysis.

Mean inter-subject correlations for each of the three dimensions (each based on 24 participants' data) were as follows: *activity* = .591, *pleasantness* = .423, *strength* = .357. It was decided to exclude participants with individual inter-subject correlations below .2 from further analyses, and, as a result, one participant was excluded from the mean *activity* rating, one (different participant) from the mean *pleasantness* rating and four from the mean *strength* rating (one of whom was the participant excluded from the *activity* analysis). To summarize, out of a total of 25 participants, the mean *activity* and *pleasantness* ratings were based on 23 participants' responses, while the mean *strength* rating was based on 20 participants' responses.

Next, in order to see how ratings of the three dimensions related to each other, inter-dimension correlations, based upon the mean rating for each of the three dimensions, were calculated. The correlation matrix for the averaged data is shown in Table 1 and reveals a moderate negative correlation between ratings of *activity* and *pleasantness*, a strong positive correlation between ratings of *activity* and *strength*, and a moderate negative correlation between ratings of *strength* and *pleasantness*.

LAG ANALYSIS

To investigate the temporal relation between the perceived emotions and the musical feature variables, a series of cross-correlation analyses were carried out. Specifically, each of the three perceived emotion ratings was cross-correlated with each of the musical feature variables. In each case, the maximal cross-correlation within the range -10 samples to $+10$ samples (-5 s to $+5$ s) indicated the lag between the musical feature and participants' response to it. Lags for each musical variable and each emotional rating are shown in Table 2. It can be seen that most lags fall between zero and four samples, although both mean pitch/strength and SD of velocity/strength are much higher. There is no lag value for articulation/strength as the maximal cross-correlation fell outside the range of accepted lags. In subsequent analyses, each variable was individually lagged according to the values in Table 2. However, because of their missing or large values, articulation, mean pitch and SD of velocity were excluded from the *strength* analysis.

REGRESSION ANALYSES

To model the experimental data, we employed ordinary least squares linear regression, in which the musical variables were used as predictors for each of the three perceived

TABLE 1 Correlation matrix showing the inter-dimension correlations between activity, pleasantness and strength

	Activity	Pleasantness	Strength
Activity	1.000	-.521	.865
Pleasantness	-.521	1.000	-.432
Strength	.865	-.432	1.000

TABLE 2 *Lags for each musical variable and each emotional rating. The lag unit is one sample (500ms), and positive values indicate the number of samples that elapsed between a musical feature occurring, and participants responding to it. There is no lag value for articulation/strength since the maximal cross-correlation fell outside the range of accepted lags*

	Activity	Pleasantness	Strength
Note density	0	0	0
Articulation	0	0	*
Mean pitch	2	3	8
SD of pitch	0	4	1
Mean velocity	1	0	1
SD of velocity	1	1	9
Tonal clarity	0	0	2
Pulse clarity	0	0	0
Dissonance	1	0	0

emotion dimensions. Some authors (e.g. Schubert and Dansmuir, 1999; Schubert, 2004) have suggested that the time-series nature of both musical and continuous response data violate a key assumption of ordinary least squares linear regression because of the non-independent nature of successive data points. This issue, known as serial correlation, was dealt with by both Schubert (2004) and Schubert and Dansmuir (1999) by differencing successive values of each variable and adding an autoregressive term in which each data point is specified to be dependent upon the one that precedes it.

We inspected our data to check for issues of serial correlation by examining the autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) of our variables. These indicated the presence of first-order serial correlation, but no autoregressive component. Thus, it was not necessary to add an autoregressive component to our regression model, but we tried differencing the data points. We did not obtain any significant models using this approach, however. We concluded that it is not possible to predict the small-scale temporal structure of the time series of this kind of data,⁷ perhaps because of the nature of the stimuli. Consequently we tried to model the coarser structure of the material by down-sampling the data, using every 12th data point in subsequent analyses.⁸ This also had the effect of reducing the presence of serial correlation in the data.

Three separate linear regression analyses were carried out, one for each of the three emotion dimensions. In each analysis, the musical variables (all nine for the *activity* and *pleasantness* analyses, and six for the *strength* model; see previous section) were entered simultaneously. Significant models emerged for *activity* [$F(9, 190) = 82.792$, $p < .001$; $R^2 = .797$; adjusted $R^2 = .787$], *pleasantness* [$F(9, 190) = 29.919$, $p < .001$; $R^2 = .586$; adjusted $R^2 = .567$], and *strength* [$F(6, 193) = 164.485$, $p < .001$; $R^2 = .836$; adjusted $R^2 = .831$]. Musical variables, and their respective beta values, for each of the three models are shown in Tables 3, 4, and 5.

It can be seen from Table 3 that ratings of perceived *activity* were positively related to note density, mean pitch, mean velocity, pulse clarity and sensory dissonance. Table 4 reveals that ratings of perceived *pleasantness* were positively related to articulation, mean pitch and tonal clarity, and negatively related to note density, SD of mean pitch, SD of velocity and pulse clarity. Table 5 shows that ratings of perceived *strength* were positively related to note density, SD of pitch, mean velocity, tonal clarity, pulse clarity and sensory dissonance.

To help visualize the success of the three models in predicting participants' ratings, the predicted values of *activity*, *pleasantness* and *strength* resulting from these models were plotted against the actual mean rating for each of these dimensions, for each of the 20 excerpts. These plots are shown in Figure 4. It can be seen that the predicted ratings of the three dimensions generally correspond quite closely to the actual mean ratings. It can also be seen, however, that the fit between predicted and

TABLE 3 *The regression model for activity. The beta values indicate the strength and direction of the relationship between each predictor variable and the mean activity rating*

Predictor variable	Beta	Sig.
Note density	.363	<.001
Articulation	-.021	NS
Mean pitch	.100	<.01
SD pitch	.035	NS
Mean velocity	.272	<.001
SD velocity	.077	NS
Tonal clarity	.051	NS
Pulse clarity	.317	<.001
Dissonance	.244	<.001

TABLE 4 *The regression model for pleasantness. The beta values indicate the strength and direction of the relationship between each predictor variable and the mean activity rating*

Predictor variable	Beta	Sig.
Note density	-.400	<.001
Articulation	.155	<.01
Mean pitch	.117	<.05
SD pitch	-.148	<.05
Mean velocity	-.076	NS
SD velocity	-.125	<.05
Tonal clarity	.220	<.001
Pulse clarity	-.196	<.001
Dissonance	-.080	NS

TABLE 5 *The regression model for strength. The beta values indicate the strength and direction of the relationship between each predictor variable and the mean activity rating*

Predictor variable	Beta	Sig.
Note density	.301	<.001
SD pitch	.119	<.001
Mean velocity	.513	<.001
Tonal clarity	.108	<.001
Pulse clarity	.177	<.001
Dissonance	.229	<.001

actual ratings varies from excerpt to excerpt and, as one would expect, because of the different amount of variance explained by the three models, between the three dimensions.

Finally, potential issues of multicollinearity were investigated in order to check the accuracy and stability of the models. An examination of two indices of multicollinearity revealed no serious concerns related to this phenomena. More specifically, mean variance inflation factors (VIFs), which indicate whether a predictor has a strong linear relationship with the other predictors, for the *activity*, *pleasantness* and *strength* models were 1.657, 1.645, and 1.624, respectively. These figures suggest that small levels of multicollinearity may be present (see Bowerman and O'Connell, 1990). However, tolerances for all variables were at least .3, and in most cases were above .6. Tolerance is the percentage of the variance in a given predictor that cannot be explained by the other predictors (values should be multiplied by 100 to obtain the percentage). Only values below .1 (.2 according to Menard, 1995) indicate serious problems (see Field, 2005), and the fact that most of the variables had high tolerances indicated no problems of multicollinearity. In sum, then, we were confident that multicollinearity was not unduly affecting the accuracy or stability of the regression models.

Discussion

Participants' ratings of *activity*, *pleasantness* and *strength* were found to relate to large-scale temporal patterns of musical features present in the improvisation excerpts. Higher *activity* ratings were best predicted by higher note density, greater pulse clarity, higher mean velocity and higher levels of sensory dissonance. Higher *pleasantness* ratings, meanwhile, were best predicted by lower note density, higher tonal clarity and lower pulse clarity. Higher *strength* ratings were best predicted by higher mean velocity, higher note density and higher dissonance. Moreover, these combinations of features accounted for between 57 and 84 percent of the variance in participants' ratings of the three dimensions. This was evidenced in the comparatively close fit between the three models' predicted ratings for the 20 excerpts and participants' actual mean ratings.

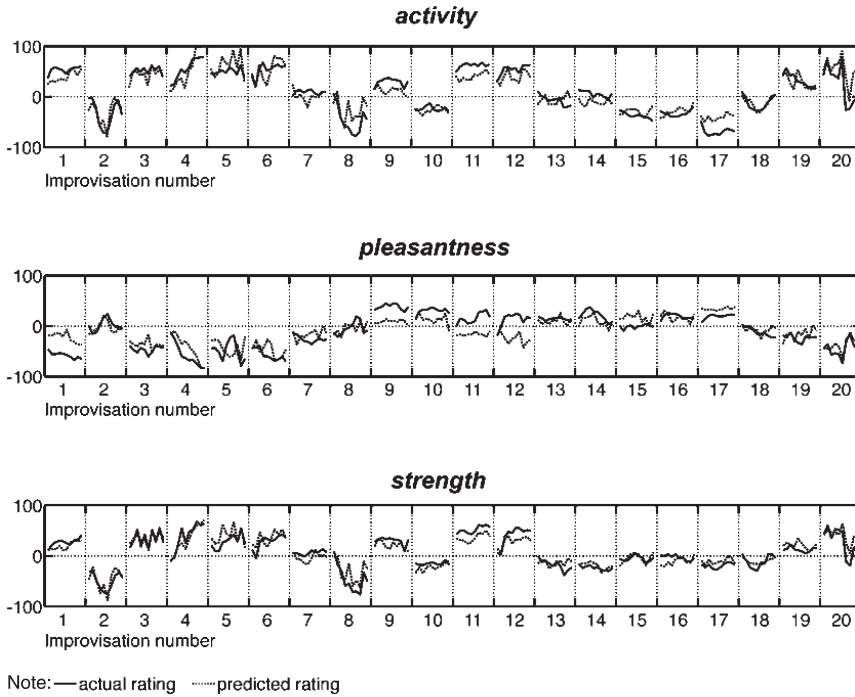


FIGURE 3 Actual and predicted ratings (vertical axes) of each of the three emotion dimensions, plotted against time (horizontal axes), for all 20 excerpts.

For each dimension, then, the combination of features that best predicted participants' ratings make rather intuitive sense: *activity* being related to a large number of notes played at a relatively high volume, often with a clear pulse; *pleasantness* being related to fewer notes played with a sense of tonality, with a less well-defined pulse; and *strength* being related to extended loud passages with lots of notes, leading to higher levels of dissonance.

Moreover, these results also largely support our tentative predictions regarding relationships between musical features and ratings along the three dimensions of emotion. Indeed, the only statistically significant relationships that did not support our hypotheses were the positive relationship between mean pitch and ratings of *pleasantness*, and the negative relationship of SD of pitch and ratings of *pleasantness*. It is difficult to explain this anomaly constructively, but perhaps future research will shed light on this finding.

Our results also compare favourably to those reported in the continuous response literature. For example, Schubert (2004) found that ratings of arousal were best explained by features such as loudness and tempo, while ratings of valence were somewhat related to melodic contour. Whilst Schubert's terminology does not map exactly onto that used in the present study, if we assume that arousal is somewhat related to activity, and valence somewhat similar to pleasantness, similarities between the two studies' findings become apparent.

For instance, higher mean velocity (which may be seen as somewhat analogous to, though not exactly the same as, Schubert's loudness), was a significant predictor of

activity ratings. As regards tempo, the present study did not examine this feature since the nature of the stimuli did not easily permit its extraction. However, tempo and note density tend to be positively correlated, so it is possible that there would have been some kind of relationship between activity and tempo in the present study.

It is a little more difficult to see how Schubert's (2004) finding that valence is related to melodic contour relates to the present study since we did not examine this musical feature (it was not easily defined in the improvisations because of their polyphonic nature and the frequent absence of clear melodic shapes). However, relationships between valence and features such as mode and articulation have been reported elsewhere (e.g. Gabrielsson and Lindström, 2001; Fabian and Schubert, 2004), and this compares well to the finding in the present study that pleasantness ratings were related to high tonal clarity and smooth articulation.

In addition to revealing a pattern of results that largely supports previous work, the present study also demonstrated a number of developments of the methodology used in similar and related studies.

The aim of the present study was to investigate temporal aspects of the relationships between musical features and emotions. To this end, we undertook a detailed description of a large set of musical improvisations and a thorough statistical analysis of the relationships between these structural descriptions and participants' continuous ratings of three emotional dimensions. The study of this complex data, the management of which required the help of computational automation, offered, in return, an informative description of these complex relationships. This study illustrated, therefore, the advantages of objective and thorough analyses of a data-rich domain of study (Clarke and Cook, 2004).

The study of the interrelation between music description and emotion offers interesting applications for both domains. On one hand, the study suggests objective characterizations of general concepts used in music therapy as a product of features directly computed from the actual description of music. On the other hand, the interrelation enables a psychological evaluation of the different methods of music description: the statistical result shows in particular the impact of each musical feature on the listeners' ratings.

Turning now to music therapy, the present study dealt with several issues highly relevant to clinical improvisation research. Because the music under analysis was, as a consequence of interactive processes between musically untrained clients and therapists, spontaneously created, listeners' ratings of its emotional content were free from biases resulting from learnt associations and familiarity. Indeed, there is a need to remove learnt and trained musical cognitive processes in order to interpret the results in a clinically relevant way, which is often the attitude of clinicians when they are interpreting client improvisations (see Pavlicevic, 1997).

Improvisation provides a particularly intensive framework for interaction (Sawyer, 2003), and this interaction is not limited only to forms of established artistic performance. One of the main reasons for the use of music in therapeutic settings is related to its communicative properties. More specifically, music is considered to share common features with verbal communication. These features are utilized to promote and enhance communication in diverse clinical conditions, an obvious example of which is aphasia (Patel, 2005). Temporal properties of music are not connected only to verbal

language but also to physical movement. Buhusi and Meck (2005) argue that interval timing is related to the coincidental activation of distinct neural cell populations. This activation obviously requires the proper function of these neural microstructures, and, conversely, the structural and functional state of these neural cell distributions most likely has an effect on the external expression of this neural timing procedure. In other words, temporal structures of musical expression and interaction reflect underlying neural mechanisms.

Thus, while clinical improvisation may not sound like 'real' music, clinicians see it as meaningful and use these elements of communication, as well as 'universal' concepts such as 'dynamic forms' or 'vitality affects' (e.g. Erkkilä, 1997; Pavlicevic, 1997) adopted from the field of developmental psychology, when describing the clinical meaning of improvisation. In terms of the present study, features such as 'density', 'velocity', 'pulse clarity' and 'dissonance' can be associated with the above-mentioned types of meaning.

It seems that at least some of the meanings of clinical improvisation can be explained by using a combination of concepts taken from traditional music analysis, psychoacoustics and mainstream psychology. Furthermore, this study supports the notion that clinical improvisation is based on meaningful psychological expression in spite of its sometimes seemingly non-musical appearance. In future work, we plan to extract a more extensive palette of musical features in order to more richly describe the musical content of the improvisations. We also plan to employ alternative methods of time-series analysis to model relationships between musical features and perceived qualities to enable the prediction of small-scale temporal structure.

ACKNOWLEDGEMENTS

The authors would like to extend their thanks to two anonymous reviewers who provided insightful and constructive feedback on earlier versions of this manuscript. The present research was supported by the Academy of Finland, project number 102253.

NOTES

1. 'Variation in pitch', to use Scherer and Oshinsky's (1977) terminology.
2. 'Loudness', to use Schubert's (2004) terminology.
3. 'Variation in loudness', to use Scherer and Oshinsky's (1977) terminology.
4. 'Loudness', to use Kleinen's (1968) terminology.
5. By musically experienced, we do not necessarily mean 'classically musically trained'. For example, a large proportion of the students in the music department at the University of Jyväskylä have received training in pop/rock, jazz and folk music styles instead.
6. These excerpts are available online at: www.jyu.fi/~luck/stimuli/improvisationexcerpts.zip
7. It should be noted that we tried to construct a single model for the whole dataset, unlike Schubert (2004) who constructed separate models for each musical stimulus.
8. While this reduced the resolution of our data, it still allowed the effective investigation of relationships between musical features and perception of emotion.

REFERENCES

- Bowerman, B.L. and O'Connell, R.T. (1990) *Linear Statistical Models: An Applied Approach* (2nd edn). Belmont, CA: Duxbury.
- Bruscia, K. (1987) *Improvisational Models of Music Therapy*. Springfield, IL: Charles C. Thomas.

- Bruscia, K. (2001) 'A Qualitative Approach to Analyzing Client Improvisations', *Music Therapy Perspectives* 19: 7–26.
- Buhusi, C.V. and Meck, W.H. (2005) 'What Makes Us Tick? Functional and Neural Mechanisms of Interval Timing', *Nature Reviews Neuroscience* 6(10): 755–65.
- Campbell, D.G. (1942) 'Basal Emotional Patterns Expressible in Music', *American Journal of Psychology* 55: 1–17.
- Clarke, E. and Cooke, N. (2004) *Empirical Musicology: Aims, Methods, Prospects*. Oxford: Oxford University Press.
- Coggiola, J.C. (2004) 'The Effect of Conceptual Enhancement in Jazz Music Selections and Jazz Experience on Musicians' Aesthetic Response', *Journal of Research in Music Education* 52(1): 29–42.
- Costa, M., Bitti, P.E.R. and Bonfiglioli, L. (2000) 'Psychological Connotations of Harmonic Musical Intervals', *Psychology of Music* 28(1): 4–22.
- Darwin, C.J., Turvey, M.T. and Crowder, R.G. (1972) 'An Auditory Analogue of the Sperling Partial Report Procedure: Evidence for Brief Auditory Storage', *Cognitive Psychology* 3: 255–67.
- Davies, J.B. (1978) *The Psychology of Music*. London: Hutchinson.
- De Backer, J. (2004) 'Music and Psychosis', unpublished master's dissertation, University of Aalborg, Denmark.
- Eerola, T. and Toiviainen, P. (2004) *MIDI Toolbox: MATLAB Tools for Music Research*. University of Jyväskylä: Kopijyvä, Jyväskylä, Finland. URL (consulted 4 May 2007): <http://www.jyu.fi/musica/miditoolbox/>
- Erkkilä, J. (1997) 'From The Unconscious to the Conscious: Musical Improvisation and Drawings as Tools in the Music Therapy of Children', *Nordic Journal of Music Therapy* 6(2): 112–120.
- Erkkilä, J. (2004) 'From Signs to Symbols, from Symbols to Words: About the Relationship Between Music and Language, Music Therapy and Psychotherapy', *Voices* 4(2). URL (consulted June 2005): <http://www.voices.no/mainissues/mi40004000150.html>
- Fabian, D. and Schubert, E. (2004) 'Expressive Devices and Perceived Musical Character in 34 Performances of Variation 7 from Bach's Goldberg Variations', *Musicae Scientiae Special Issue 2003–2004*: 49–71.
- Field, A. (2005) *Discovering Statistics Using SPSS*. London: Sage.
- Fraisse, P. (1982) 'Rhythm and Tempo', in D. Deutsch (ed.) *Psychology of Music*, pp. 149–80. New York: Academic.
- Fredrickson, W.E. (1995) 'A Comparison of Perceived Musical Tension and Aesthetic Response', *Psychology of Music* 23(1): 81–7.
- Fredrickson, W.E. (1999) 'Effect of Musical Performance on Perception of Tension in Gustav Holst's First Suite in E-flat', *Journal of Research in Music Education* 47(1): 44–52.
- Fredrickson, W.E. and Coggiola, J.C. (2003) 'A Comparison of Music Majors' and Nonmajors' Perceptions of Tension for Two Selections of Jazz Music', *Journal of Research in Music Education* 51(3): 259–70.
- Gabrielsson, A. and Lindström, E. (2001) 'The Influence of Musical Structure on Emotional Expression', in P.N. Juslin and J.A. Sloboda (eds) *Music and Emotion: Theory and Research*, pp. 223–48. Oxford: Oxford University Press.
- Goins, W.E. (1998) 'The Effect of Mood States: Continuous Versus Summative Responses', *Journal of Music Therapy* 35(4): 242–58.
- Hampton, P.J. (1945) 'The Emotional Element in Music', *Journal of General Psychology* 33: 237–50.
- Helmholtz, H. (1877) *On the Sensations of Tone*, trans. A.J. Ellis. New York: Dover.
- Juslin, P.N. and Sloboda, J.A. (2001) *Music and Emotion: Theory and Research*. London: Oxford University Press.

- Kameoka, A. and Kuriyagawa, M. (1969) 'Consonance Theory', *Journal of the Acoustic Society of America* 45(6): 1452–69.
- Kleinen, G. (1968) *Experimentelle Studien zum Musikalischen Ausdruck* [Experimental Studies of Musical Expression], Hamburg: Universität Hamburg.
- Krumhansl, C.L. (1990) *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.
- Krumhansl, C.L. (1996) 'A Perceptual Analysis of Mozart's Piano Sonata k.282: Segmentation, Tension, and Musical Ideas', *Music Perception* 13(3): 401–23.
- Krumhansl, C.L. (1997) 'An Exploratory Study of Musical Emotions and Psychophysiology', *Canadian Journal of Experimental Psychology* 51(4): 336–52.
- Krumhansl, C.L. (1998) 'Topic In Music: An Empirical Study of Memorability, Openness, and Emotion in Mozart's String Quintet in C Major, and Beethoven's String Quartet in A minor', *Music Perception* 16(1): 119–34.
- Krumhansl, C.L. and Schenck, D.L. (1997) 'Can Dance Reflect the Structural and Expressive Qualities of Music? A Perceptual Experiment on Balanchine's Choreography of Mozart's Divertimento No. 15', *Musicae Scientiae* 1(1): 63–85.
- Lee, C. (2000) 'A Method of Analysing Improvisations in Music Therapy', *Journal of Music Therapy* 37(2): 147–67.
- Madsen, C.K. (1997) 'Emotional Response to Music as Measured by the Two-Dimensional CRDI', *Journal of Music Therapy* 34(3): 187–99.
- Madsen, C.K. and Fredrickson, W.E. (1993) 'The Experience of Musical Tension: A Replication of Nielsen's Research using the Continuous Response Digital Interface', *Journal of Music Therapy* 30: 46–63.
- Menard, S. (1995) *Applied Logistic Regression Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07–106. Thousand Oaks, CA: Sage.
- Nordoff, P. and Robbins, C. (1965) *Music Therapy for Handicapped Children*. New York: Rudolf Steiner Publications.
- Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. (1957) *The Measurement of Meaning*. Champaign, IL: University of Illinois Press.
- Patel, A.D. (2005) 'The Relationship of Music to the Melody of Speech and to Syntactic Processing Disorders in Aphasia', *Annals of the New York Academy of Sciences* 1060: 59–70.
- Pavlicevic, M. (1997) *Music Therapy in Context: Music, Meaning and Relationship*. London: Jessica Kingsley.
- Pavlicevic, M. (2000) 'Improvisation in Music Therapy: Human Communication in Sound', *Journal of Music Therapy* 37(4): 269–85.
- Plomp, R. and Levelt, W.J.M. (1965) 'Tonal Consonance and Critical Bandwidth', *Journal of the Acoustic Society of America* 38: 548–60.
- Sawyer, R.K. (2003) *Group Creativity: Music, Theatre, Collaboration*. Mahwah, NJ: Lawrence Erlbaum.
- Scherer, K.R. and Oshinsky, J.S. (1977) 'Cue Utilization in Emotion Attribution from Auditory Stimuli', *Motivation and Emotion* 1: 331–46.
- Schubert, E. (1999) 'Measuring Emotion Continuously: Validity and Reliability of the Two-Dimensional Emotion-Space', *Australian Journal of Psychology*. Special Issue: *Music as a Brain and Behavioural System* 51(3): 154–65.
- Schubert, E. (2001) 'Continuous Measurement of Self-Report Emotional Response to Music', in P.N. Juslin and J.A. Sloboda (eds) *Music and Emotion: Theory and Research*, pp. 393–414. Oxford: Oxford University Press.
- Schubert, E. (2004) 'Modelling Perceived Emotion with Continuous Musical Features', *Music Perception* 21(4): 561–85.

- Schubert, E. and Dansmuir, W. (1999) 'Regression Modelling Continuous Data in Music Psychology', in S.W. Yi (ed.) *Music, Mind, and Science*, pp. 298–352. Seoul: Seoul National University Press.
- Smeijsters, H. (1997) *Multiple Perspectives: A Guide to Qualitative Research in Music Therapy*. Gilsum, NH: Barcelona Publishers.
- Swanwick, K. (1973) 'Musical Cognition and Aesthetic Response', *Psychology of Music* 1(2): 7–13.
- Toiviainen, P. (2001) 'Real-Time Recognition of Improvisations with Adaptive Oscillators and a Recursive Bayesian Classifier', *Journal of New Music Research* 30: 137–48.
- Toiviainen, P. and Krumhansl, C.L. (2003) 'Measuring and Modelling Real-Time Responses to Music: The Dynamics of Tonality Induction', *Perception* 32(6): 741–66.
- Toiviainen, P. and Snyder, J.S. (2003) 'Tapping to Bach: Resonance-based Modelling of Pulse', *Music Perception* 21(1): 43–80.
- Treisman, A.M. (1964) 'Verbal Cues, Language, and Meaning in Selective Attention', *American Journal of Psychology* 77: 206–19.
- Van Noorden, L. and Moelants, D. (1999) 'Resonance in the Perception of Musical Pulse', *Journal of New Music Research* 28: 43–66.
- Vink, A. (2001) 'Living Apart Together: A Relationship Between Music Psychology and Music Therapy', *Nordic Journal of Music Therapy* 10(2): 144–58.
- Warren, R.M. (1993) 'Perception of Acoustic Sequences: Global Integration Versus Temporal Resolution', in S. McAdams and E. Bigand (eds) *Thinking in Sound*, pp. 37–68. Oxford: Oxford University Press.
- Wedin, L. (1972) 'Multidimensional Study of Perceptual-Emotional Qualities in Music', *Scandinavian Journal of Psychology* 13: 241–57.
- Westergaard, P. (1975) *An Introduction to Tonal Theory*. New York: W.W. Norton.
- Wigram, T. (2004) *Improvisation: Methods and Techniques for Music Therapy Clinicians, Educators and Students*. London: Jessica Kingsley.
- Wosch, T. (2003) 'NEoM: A Research Network on Emotion in Music Therapy', in D. Aldridge and J. Fachner (eds) *Music Therapy World*, 10th World Congress of Music Therapy, Oxford, UK, 2002, Info-CD ROM V. Witten: University of Witten-Herdecke.

GEOFF LUCK is an Associate Professor in the Department of Music at the University of Jyväskylä. His research interests include the characteristics of temporal and expressive conducting gestures, between-person coordination, and the development of synchronization abilities across the life span.

Address: Department of Music, PO Box 35 (M), 40014, University of Jyväskylä, Finland
[email: luck@campus.jyu.fi]

PETRI TOIVIAINEN is a specialist in computational modelling of music perception and cognition. Currently he holds a permanent post as a professor of cognitive musicology at the University of Jyväskylä. Before this, he has been, among other things, visiting professor at the Department of Psychology of Cornell University (USA) and visiting researcher at the Institute for Psychoacoustics of Ghent University (Belgium). He is the head of the Finnish Centre of Excellence in Interdisciplinary Music Research.

Address: Department of Music, PO Box 35 (M), 40014, University of Jyväskylä, Finland.
[email: ptoiviai@campus.jyu.fi]

JAAKKO ERKKILÄ is an expert in music psychotherapy where his special concerns are theory of psychodynamic music therapy, clinical improvisation, and qualitative research of clinical processes. He has published papers, books and given several talks at international congresses. Currently he holds a permanent post as a professor of music therapy at the University of Jyväskylä. He is the director of Music Therapy Masters Programme (University of Jyväskylä) and the director of Music Therapy Clinical Training (Eino Roiha Institute). Currently he acts as the Vice-President of the European Music Therapy Confederation.

Address: Department of Music, PO Box 35 (M), 40014, University of Jyväskylä, Finland.
[email: jerkkila@campus.jyu.fi]

OLIVIER LARTILLOT is post-doctoral researcher at the Department of Music of the University of Jyväskylä, in the fields of computational musicology and music information retrieval. He works on the NEST project, Tuning the Brain for Music, supported by the European Commission (028570), and serves as reviewer for several international journals.

Address: Department of Music, PO Box 35 (M), 40014, University of Jyväskylä, Finland.
[email: lartillo@campus.jyu.fi]

KARI RIIKKILÄ is a music therapist and obtained his MA in music therapy in 2000. In addition to clinical work and lecturing, he is a researcher at the University of Jyväskylä, and works on the Academy of Finland research project, Intelligent Music Systems in Music Therapy.

Address: Department of Music, PO Box 35 (M), 40014, University of Jyväskylä, Finland.
[email: katariik@campus.jyu.fi]

ARTO MÄKELÄ is a music therapist at Satakunta District of Services for the Intellectually Disabled, Finland.

[email: arto.makela@sataehp.fi]

KIMMO PYHÄLUOTO is a music therapist for Pääjärvi Federation of Municipalities, Finland.

[email: pyhaluoto@gmail.com]

HEIKKI RAINE is a music therapist at the Rinnekoti-Foundation, Finland.

[email: heikki.raine@rinnekoti.fi]

LEILA VARKILA is a music therapist for Pääjärvi Federation of Municipalities, Finland.

[email: leila.varkila@paajarvenky.fi]

JUKKA VÄRRI is a music therapist for Suojarinne Federation of Municipalities, Finland.

[email: jukkavarri@gmail.com]