# Playing with Affect: Music Performance with Awareness of Score and Audience


**Steven R. Livingstone**

School of Information Technology and Electrical Engineering,
The University of Queensland
And
Australian CRC for Interaction Design, QUT
srl@itee.uq.edu.au

**Ralf Muhlberger**

School of Information Technology and Electrical Engineering,
The University of Queensland
ralf@itee.uq.edu.au

**Andrew Brown**

School of Music
Queensland University of Technology
And
Australian CRC for Interaction Design, QUT
a.brown@qut.edu.au



## Abstract

An exquisite music performance can move an audience to tears of joy, or those of sorrow. The job of a good performer then is to convey not just the music's notated structure, but also its emotional metadata. In time, the performer will also learn to respond musically to the state of a live audience; matching their emotional ebb and flow to maintain interest for the concert's entirety. This work will discuss an Affective Performance framework in which compositions are marked up with emotional intent, or narrative. This mark-up directs the emotive adaptation of the symbolic score's reproduction, enhancing the computer music's realism. Examining the cognitive model of emotions appraisal theory, we highlight some key evidence underlying the principle of music expectancy and its central role in maintaining a listener's musical interest. This knowledge will play a significant role in developing an adaptive music engine in which the traditionally static score is manipulated in real time. We will also show how new Affective Computing technologies for reading the emotional state of users can be employed to further adapt the performance to account for audience emotional state. We end with an illustration of the framework, with examples in computer gaming engines and enhanced live performances for a distributed audience.


## 1 Introduction

The most common criticism levelled at computer generated music is its lack of realism. While sound reproduction has improved with instrument sample quality, the underlying process of converting a work's structural information into an emotive, humanistic performance has not. However, in order to produce an acceptable level of realism we must go beyond the simple reproduction of symbolic notation. To do this, a framework is required which can operate on the level of music awareness. By examining what in music instils emotions in listeners, we can employ this knowledge in a computing aspect. Marking up the traditionally static music score with emotional performance metadata will give symbolic music a degree of "humanity". To gather this understanding, a multidisciplinary approach to music and emotion is undertaken. How this music is then performed in a live environment is also an important issue. Traditionally music is thought of as a simple three step process: composer → performer → listener, however this is a model taken from the Western classical music scene in which the listener sits quietly throughout a performance. In the context of modern live performances and computer music, this is no longer the case. Crucially, the concept of listener feedback is a powerful affective medium with which modern performers attune; the same feedback process is possible in the computer music environment. In order to do this, the interplay of feedback in the music environment, along with listener attitudes is formalised into an information model.

The structure of the paper is as follows; in section 2 we will outline the Affective Performance framework. This will be followed by a discussion of listener attitudes in section 3. An examination of the concept music interest, along with existing music-emotion research is discussed in section 4. In section 5 we will give an outline of the music environment and listener attitude model, along with a discussion of how music aesthetics are incorporated. We conclude in section 6 with an overview of the material presented in this paper.


Livingstone, S. R., Muhlberger, R. and Brown, A. R. 2005. "Playing With Affect: Music performance with awareness of score and audience." In T. Opie and A. R. Brown (eds.) Australasian Computer Music Conference. Brisbane: ACMA, pp. 92-98.

## 2 Framework

To understand the context in which the Affective Performance framework will operate, we refer the reader to figure 1. Each of the models listed represents a different scenario of music score reproduction. In each, either a human or computer can take the place of a performer. Each model is discussed in the following subsections.

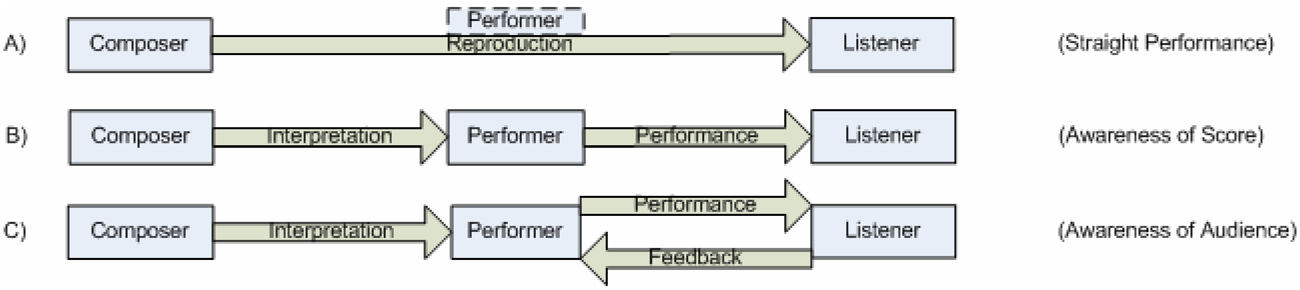


**Figure 1** – Performance Models

### 2.1 Straight Performance

The model of straight performance is the earliest and most basic model of score reproduction. In this form, the performer (computer/human) simply reproduces the music score exactly as notated. Early examples of this model were present in the Baroque era, where the keyboard works of J. S. Bach were written for the harpsichord, a forerunner to today's pianoforte. Here, the performer was incapable of personalising the score, and acted simply as a means to reproduce the notation. A similar parallel is found today in symbolic computing music. Lacking any form of intra-note dynamics, the majority of midi performance is flat and emotionless; an undesirable model.

### 2.2 Awareness of Score

In this model the performer plays an active role in score reproduction, with a variety of personalising characteristics. This model forms the majority of today's classical music performances and pre-recorded computer game music. The humanising element allows for reproduction in a way that the composer originally intended, with a great degree of emotion transmitted in the performance/interpretation process. To enable symbolic computer music to appear realistic, it must be marked-up with emotional metadata similar to that used by a human performer. This mark-up will be discussed later in section 4.

### 2.3 Awareness of Audience

The final model is the optimum scenario, and the latest to appear in Western music. In this model, the audience takes an active role in influencing the reproduction process of the performer. This form is common in today's popular music scene, where audience members cheer and sing along with the performer and work to control the 'energy flow' of the concert. Both the performer and audience work in concert to produce a tailored music experience. This is the idealised scenario in computer gaming in which the music is tailored to the emotional state of the user. In the context of distributed human performance, this feedback model can be applied to assist in live, augmented web-cast performances. Here, emotional metrics can be gathered and summarised by a computational architecture and fed back to the performer in order to tailor their response to match the desire of the audience, or to filter their performance individually for each listener tuning in. In essence, the aim is to bring the audience back into the loop of modern music performance.

### 2.4 Architecture

To achieve the goals outlined above for an affective mark-up system with audience feedback, our framework requires four components.

i. State Database – Store a listener's, audience and performer's attitudes and state information for feedback and tailoring (see section 3)
ii. Audience Awareness – A set of affective computing technologies and methodologies for capturing emotional state and attitudes
iii. Interpretive Engine – This component forms the core of the framework, providing the emotive processing and mark up
iv. Audio Production – Handles the production of sound from the emotively marked-up music

A discussion of the architecture's implementation is delayed until section 5.1.

## 3 User Attitudes

At its heart music is about making the listener feel something. Through a narrative of sound the composer is capable of weaving an abstract emotional pathway with which to guide the listener. However, the process of communicating emotion in the real world is far more complex than the commonly idealised situation. The most commonly overlooked factor in the music life cycle is the production environment. While a typical environment consists of an audience in an enclosed space listening to a pre-ordained work, this is a static, uni-directional view that rejects the ability of the listener to modify both a performer's style of playing and the work itself in the case of real-time generated music or performer ad libitum.

An important consideration when selecting music for a target audience is the notion of user attitudes. However, the interplay of attitudes in the music environment has been largely ignored by the music research community. Attitudes are a powerful cognitive tool which enables us to quickly categorise, influence and reason about new stimuli. They can range from the flexible, such as "what colour to paint the new room", to those of obstinate rigidity, "what political party to vote for". Attitudes are believed to be composed of three distinct elements; an affective component for negative/positive emotions about the object, a cognitive component for beliefs and perceptions of the object, and a behavioural component for user response tendencies (Petty, Fabrigar et al. 2003). To take an example: a group of young music goers gather to hear their favourite death-metal band play; at the moment their anticipated arrival, a classical chamber orchestra walks out on the stage instead and begins to play some Mozart. While an audience of classical enthusiasts would be enraptured by these sounds, our concert goers probably harbour negative feelings towards classical music; the listeners may perceive such music as boring, while behaviourally they may stop listening and trying to prevent the performers from continuing by booing. Thus music attitudes must be considered when composing music for a target audience. Clearly though, when music is pre-selected either for or by an audience, the genre of music chosen will generally reflect their primary preferences/attitudes. The notion of attitudes in our model is one of general sorting, or fine-tuning to improve the listening experience.

An appropriate feedback model must also incorporate the current internal state of both the listener and the performer. A person's internal state can be indicated by three measures: emotions, mood and physical. Emotions are the feelings which exist on the order of seconds (eg happy, sad, angry); mood represents those more enduring states which exist on the order of hours or days (eg depressed, worried, in love) (Levenson 2003). Lastly, physical state is included to denote level of alertness (eg sleepy, energised); here properties such as illness or injury are not considered as their effects generally fall under the category of mood. The feedback model with incorporated user attitudes is examined in section 5.

## 4 Music Interest at the Cognitive Level

The goal of any music affect system is the capacity to influence the emotions of its listeners. The question is of course, how? While at first appearing an insurmountable problem with layers of inter-acting factors, the issue can be tackled with a straight-forward approach by examining existing music-emotion research. Since the turn of the 20$^{th}$ century, a growing bank of music-emotion rules has been derived from hundreds of empirical studies. The literature refers to rule types that typically fall under one of two umbrellas: the structural or the performance. Structural aspects refer to those which are inherent in the music's notated score, such as tempo, dynamics, and melodic direction. Performance rules relate to all measures used by a performer in translating the static score into audible information, such as rubato, intonation, and vibrato.

Recent comprehensive reviews of both the structural (Schubert 1999; Gabrielsson and Lindstrom 2001; Gabrielsson and Juslin 2003) and the performance (Gabrielsson 1999; Juslin 2001; Gabrielsson 2003) aspects has greatly expedited research proceedings. In the majority of music-emotion rule studies, listeners were asked to rate a music work's aesthetic qualities by selecting responses from a checklist of adjectives. The most commonly used checklist is Hevner's, consisting of sixty-seven words arranged into eight emotion clusters (Hevner 1936). More recently, the use of 2-dimensional space for modelling emotion has proven popular (Russell 1989). This is a consolidation of the original three-dimensional space proposed by Wundt in the late 19$^{th}$ century (Scherer 2004). The 2-dimensional model is capable of representing the majority of measurable emotions, with the horizontal 'valence' axis, for the positive-negative, and the vertical 'activation' axis (or arousal), for excited-calm. This model has been used successfully in user response music testing (Schubert 1999); its continuous numerical representation of emotions makes it a prime candidate for computational implementation, as used by the author in accompanying work. However, reducing emotional representation to two dimensions is a not a lossless transformation; certain emotions cannot be easily described as just a function of activation and valence. Examples of terms not easily described include annoyed, majestic, frustrated, sentimental and sacred. A second problem is that terms which fall close together on the valence-activation scales, such as anger and fear, are markedly dissimilar. To date however, a 2-dimensional continuous model of emotions has proven the most effective, and operates at the heart of emotional mark-up in our Affective Performance framework.

Of the hundreds of studies carried out on empirical music psychology, little if any have investigated the more complex phenomenon of music interest. While music rules are an effective indicator of the emotion of the work, these are just the tools that a composer employs. Music rules alone do not explain how they can be used as a gestalt to capture the heart of a listener for 30 minutes. Leonard Meyer's treatise on music and emotion in 1956 redefined the study of music and emotion, opening up the field to scientific critique. His theory of music expectancy offered a solid footing on which to examine the affective qualities of music structure. The underlying premise that "*the mind, governed by the law of Prägnanz, is continually striving for completeness, stability and rest*" (Meyer 1956), is based on a psychological theory of pattern matching in which we are constantly seeking identifiable patterns in order to reason about them. In music, the listener experiences a tension-relaxation flow through the successful completion or violation of musical expectation. Since its inception, the study of music expectation has received considerable attention (Eerola 2003); and has formed the basis of extended theories, such as Narmour's Implication-Realisation model (Narmour 1999) and Lerdahl's Harmonic-Tension model (Lerdahl 2001). Strangely, despite the level of research of both music emotions and music expectancy, the two have yet to be reconciled into a cohesive theory of music interest and emotion.

The problem has been that expectation is a cognitive phenomenon and as such requires a cognitive model of emotions (Krumhansl 2002). Appraisal theory is one such model (Ellsworth and Scherer 2003; Ortony 2003); it describes environmental stimulus as being assessed against a set of dimensions in order to determine its significance to the user; the result of which is the generation of an appropriate physiological and emotional response to best deal with the stimulus. In our context of music, the core appraisal dimension is that of music expectancy; where our goal is the closure or expected destination of music fragments. We propose that music tension acts at the level of cognitive appraisal, and that it is the flow and control of tension that generates and sustains listener interest. This notion fits with anecdotal evidence, where the intensity of listener emotion is greater when the listener is cognitively focused on the work, than when the music is simply played in the background. Support for music interest as a function of music tension/expectancy, and the integration of music expectancy into a model of music appraisal theory finds support from three sources. Recent work examining how listeners remember affective works found a strong correlation between the affect intensity gradient and music enjoyment/music recall (Rozin, Rozin et al. 2004). It was concluded that the most memorable or enjoyable works were those with large affect gradients, where the music moved between affective peaks and valleys. Thus, a theory of music enjoyment/interest would require increased levels of tension to be strongly correlated with all forms of increased affect (happy, sad, fearful etc). In (Krumhansl 2002) this is was observed, where it was found that tension was strongly correlated with all three basic emotions happiness, sadness and fear. However, too much tension can be a bad thing, if the music is always moving forward with little discernable pattern music interest will drop off (a common complaint against modern classical music). As Meyer puts it "*Completion is possible only where there is shape and pattern*" (Meyer 1956). Therefore, for a music work to maintain interest the music must strike a balance between too little and too much change, or a balance of original music material. One form of completion outlined by Meyer is that of melodic completeness. Referring to the work of (Simonton 2001), we point to the graph of melodic originality as a function of effective music enjoyment/popularity. The work found that optimum level music enjoyment occurs at a level of mid-range melodic originality, a finding strongly inline with our requirement for music tension.

## 5 Music Environment Representation

In figure 2 we incorporate the notion of user attitudes and internal state (see section 3) into the music environment. Using an entity-relational model, we have broadened the music environment to incorporate all elements in the music process, from compositional roots to live reproduction. Referring to figure 2, attitudes are connected to the three human elements: performer, audience and listener. From the position of the performer and listener, attitudes regarding the various entities will affect how the music is portrayed and received emotionally. The inclusion of audience attitudes is also a powerful component in the live performance feedback scenario, as group behaviours and collective opinion can affect both individual listener attitudes and those of the performer. The following scenario illustrates the power a feedback model allows.

> A group of listeners gather to hear a lively performance from their favourite popular band. After about 20 minutes of upbeat playing, the audience begins to show signs of slowing. The change in audience behavioural motions (less jumping, more yawning etc) may also affect

the attitudes of any listeners who have not slowed down, causing them to desire a slower pace. Sensing this, the performer reduces the speed of the performance to respond to the audience's emotional state.

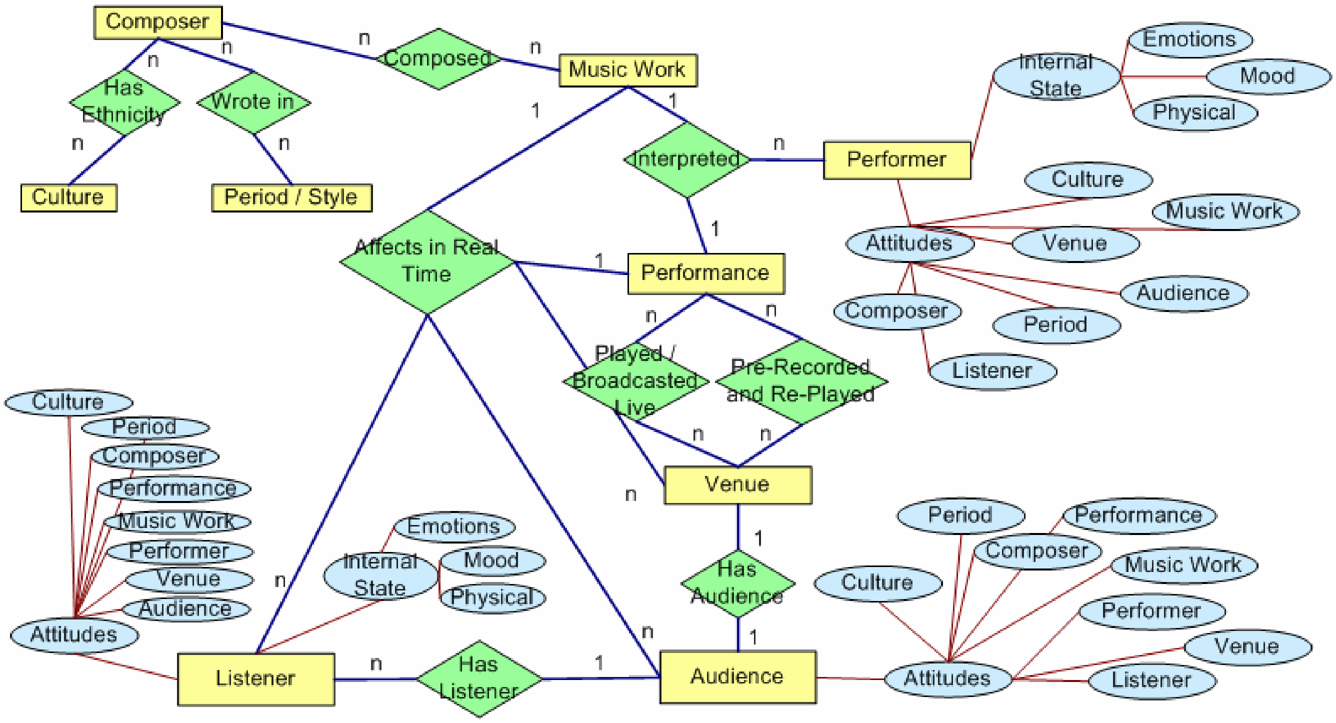


**Figure 2** – Music Environment

Formalising the music environment affords the Affective Performance framework a greater degree of control over the music process by specifically permitting user feedback. By incorporating both the user and audience into the environment, the music process ceases to be a uni-directional one. Through handing a degree of control to the listener, an increase in musical enjoyment is expected.

As seen in section 4, to accurately reflect the emotion and interest level of a musical work in our framework we must return to the original three-dimensional emotional representation as proposed by Wundt, incorporating the tension/relaxation axis into the existing activation-valence space. Through the use of music rules compiled from the literature above we now have an adequate emotional representation model for our Affective Performance framework in the formalised music environment.

### 5.1 Framework Implementation

As outlined in section 2.4 the Affective Framework operates with four main components. Component one is the state database, used to store and process listener attitudes (see section 3) and emotional state information it acts as an information repository for component 2 input, while providing data to component 3. Knowledge of a listener's attitudes/preferences with regards to genres, performers and performances can be used to build a detailed preference library which can be used to individually tailor the music experience. To gather the set of user attitudes and audience emotional responses for deductive reasoning, the system requires a set of affective computing technologies and methodologies. These technologies form component two of the Affective Performance framework. Feedback metrics such as mouse movement behaviours, keystroke and blinking rates can be used to determine the relative emotional state of the user; the technology and knowledge of which comes from the burgeoning field of affective computing (Picard, Papert et al. 2004). Component three forms the core of our framework, implementing an interpretive music-emotion engine which marks up the symbolic music score with a performance narrative. This mark-up will invest a sense of realism, where the generated music will more closely resemble the composer's original intentions (see section 4). This component will also be concerned with music interest, adapting the pre-composed music score to meet the changing environment offered in computer game play. The fourth and final component of the framework is concerned with generating sound from the emotionally marked-up music. Utilising the dynamic music generation engine AiME, the framework implements a series of filters and rules to modify the symbolic music in real time. Operating at this level of granularity allows for essentially an 'instantaneous' reflection of environmental/game events within the music, while still allowing for seamless integration into the overall score.

In the future, we will move to the testing phase of music interest and proceed with the implementation of music-emotion rules. User emotion response tools and methodologies have already been developed and used to gather preliminary results for component's 3 and 4 of the framework (see section 2.4).

## 6 Conclusions

In this paper we have outlined the Affective Performance framework, a system of listener attitudes and affective mark-up for symbolic computer music. With an integrated model of listener attitudes and feedback, the system will add a new level of realism to computer music through awareness of intended emotional response to music and audience feedback. The framework is designed to integrate with computer gaming environments and to assist live music performances in a distributed context.

## References


- Eerola, T. (2003). The Dynamics of Musical Expectancy, UNIVERSITY OF JYVÄSKYLÄ.
- Ellsworth, P. C. and K. R. Scherer (2003). Appraisal Processes in Emotion. Handbook of Affective Sciences. Davidson, Scherer and Goldsmith.
- Gabrielsson, A. (1999). The Performance of Music. The Psychology of Music. D. Deutsch.
- Gabrielsson, A. (2003). "Music performance research at the millennium." The Psychology of Music.
- Gabrielsson, A. and P. N. Juslin (2003). Emotional Expression in Music. Handbook of Affective Sciences, Oxford University Press.
- Gabrielsson, A. and E. Lindstrom (2001). The Influence of Musical Structure on Emotional Expression. Music and Emotion, theory and research.
- Hevner, K. (1936). "Experimental studies of the elements of expression in music." American Journal of Psychology **48**: 246-268.
- Juslin, P. N. (2001). Communicating emotion in music performance: a review and a theoretical framework. Music and Emotion, theory and research.
- Krumhansl, C. L. (2002). "Music: A Link Between Cognition and Emotion." Current Directions in Psychological Science **11**(2).
- Lerdahl, F. (2001). Tonal Pitch SPace, Oxford University Press.
- Levenson, R. W. (2003). Autonomic Specificity and Emotion. Handbook of Affective Sciences. Davidson, Scherer and Goldsmith.
- Meyer, L. B. (1956). Emotion and Meaning in Music, The University of Chicago Press.
- Narmour, E. (1999). Hierarchical Expectation and Musical Style. The Psychology of Music. D. Deutsch, Academic Press.
- Ortony, A. (2003). On making believable emotional agents believable. Emotions in Humans and Artifacts. R. Trappl, P. Petta and S. Payr.
- Petty, R. E., L. R. Fabrigar, et al. (2003). Emotional Factors in Attitudes and Persuasion. Handbook of Affective Sciences. Davidson, Scherer and Goldsmith.
- Picard, R. W., S. Papert, et al. (2004). "Affective Learning - A manifesto." BT Technology Journal 22(4).
- Rozin, A., P. Rozin, et al. (2004). "The Feeling of Music Past: How Listeners Remember Musical Affect." Music Percpetion **22**(1).
- Russell, J. A. (1989). Measures of emotion. Emotion: Theory research and experience. R. P. H. Kellerman, New York: Academic Press. **4:** 81-111.
- Scherer, K. R. (2004). "Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them?" Journal of New Music Research **33**(3): 239-251.
- Schubert, E. (1999). Measurement and Time Series Analysis of Emotion in Music, University of New South Wales. PhD Thesis.
- Simonton, D. K. (2001). Emotion and Composition in Classical Music: Historiometric Perspectives. Music and Emotion, theory and research.