UPPSALA UNIVERSITET

*Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 36

# A Novel Approach to Teaching Emotional Expression in Music Performance

JESSIKA KARLSSON

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA
2008

ISSN 1652-9030
ISBN 978-91-554-7089-0
urn:nbn:se:uu:diva-8456

Dissertation presented at Uppsala University to be publicly examined in Sal IV, Universitetshuset, Uppsala, Friday, March 14, 2008 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

**Abstract**

Karlsson, J. 2008. A Novel Approach to Teaching Emotional Expression in Music Performance. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 36. 57 pp. Uppsala. ISBN 978-91-554-7089-0.

One of the most important aspects of music performance is the expression of emotions, yet research has suggested that this skill is neglected in music education. The aim of this thesis was thus to develop and test a novel and empirically-based approach to teaching emotional expression in music performance.

Study I explored the nature of instrumental teaching in its natural context, with a focus on emotional expression. Although there were individual differences among teachers, a common feature was a lack of clear goals, specific tasks, systematic teaching patterns, and informative feedback.

Study II presented and tested a computer program that analyzes music performances and offers informative feedback, including specific suggestions on how to enhance the emotional expression. Performers were randomly assigned to one of three conditions: (1) feedback from the computer program, (2) feedback from music teachers, and (3) repetition without feedback. The results indicated the greatest improvements in communication accuracy for the computer feedback group, but although the computer program was rated as easy to understand and use, performers did not want to use it in the future.

Study III explored whether the negative views towards the computer program were due to negative attitudes towards computers or a dislike of the characteristics of the actual feedback contents. Results from a deception experiment revealed that the mere belief that the feedback derived from a teacher yielded higher quality ratings, but so did also feedback that did indeed derive from a teacher. The latter feedback was perceived as more detailed.

The thesis shows that it is possible for performers to improve their abilities to express emotions through computer-assisted teaching, but suggests that the feedback may benefit from including human-like aspects such as encouragement, examples, and explanations in order to make it more attractive in the eyes of its potential users.

*Keywords:* music performance, emotion, expression, feedback, computer-based teaching, music education

*Jessika Karlsson, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden*



ISSN 1652-9030
ISBN 978-91-554-7089-0
urn:nbn:se:uu:diva-8456 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8456)

# List of papers

This thesis is based on the following studies, which will be referred to in the text by their Roman numerals:

I Karlsson, J., & Juslin, P. N. (in press). Musical Expression: An Observational Study of Instrumental Teaching. *Psychology of Music*.

II Juslin P. N., Karlsson, J., Lindström, E., Friberg, A., & Schoonderwaldt, E. (2006). Play It Again With Feeling: Computer Feedback in Musical Communication of Emotions. *Journal of Experimental Psychology: Applied, 12*, 79-95.

III Karlsson, J., Liljeström, S., & Juslin, P. N. (in press). Teaching Musical Expression: Effects of Production and Delivery of Feedback by Teacher vs. Computer on Rated Feedback Quality. *Music Education Research.*

# Contents

# Introduction

> "I see a lot of technical musicians, and very few of them seem to have that feel that goes along with it. They're able to do technically a lot more than the next guy, but for some reason it doesn't communicate…" Musician cited in Boyd and George-Warren (1992, pp. 104-105)

> "Ah, expression! It's what the majority of musicians believe music to be. Yet, for some unknown reason, it is rarely talked about." Music teacher cited in Vosskuhler (2005, p. 1)

One of the most important aspects of music performance is the ability to play expressively (e.g., Juslin & Laukka, 2004; Laukka, 2004; Lindström, Juslin, Bresin, & Williamon, 2003) and thus to 'move' listeners (Juslin & Västfjäll, in press). An expressive performance is often what makes people prefer one musician over another, and it is expression that makes new interpretations of familiar music pieces possible. Most performers and listeners define musical expression in terms of communicating emotions (e.g., Lindström et al., 2003; Laukka, 2004). Thus, to master the skill of emotional expression in music is an important goal for a performer. Given the importance of expression in music performance, it is reasonable to expect that music teachers devote a lot of their time to developing this skill. However, on the contrary, there is some evidence that expression is neglected in music education (see Persson, 1993; Rostvall & West, 2001; Tait, 1992), perhaps because it is often regarded as a skill that reflects talent and thus cannot be learnt (Sloboda, 1996), or because knowledge about expression is mostly tacit and therefore difficult to convey in words (Hoffren, 1964).

The present thesis is part of a project, *Feedback-learning of Musical Expressivity* (*Feel-ME*), which aimed to define the nature of expression in music performance, and to develop new methods for teaching expression, particularly expression of emotions. The focus of this thesis is on the latter aspect of the project: Is it possible to develop a computer program that can improve a performer's ability to express emotions? The thesis is based on three empirical studies. In Study I, the nature of instrumental teaching is explored in detail to elucidate possible problems with the current teaching. Study II presents and evaluates a new computer-assisted teaching method aiming to enhance music performers' emotional expression. In Study III, various factors that may influence performers' views on computer-assisted teaching of expression are explored.

The thesis is organized in the following way. Firstly, a background to the problems surrounding musical expression in music education and in earlier research is outlined. Secondly, a new empirically-based approach to teaching expression is presented. Thirdly, the specific aims of the thesis are stated. Fourthly, the three empirical studies are summarized. Finally, the findings from the studies are discussed.

# Traditional approaches to expression

## Expression in music education

### The tendency to neglect expression

Though expressive skills are important in music performance, there is some evidence that teaching tends to focus on other aspects. However, real-world investigations of instrumental teaching are quite rare. Most early research was carried out in laboratories and focused on the personality characteristics of teachers, whereas subsequent studies have focused on teachers' behaviors (e.g., teacher activities, verbal instructions, gestures). Generally, research suggests that most instrumental teaching focuses on the reproduction of specific works within the Western art music tradition (Hallam, 1997), and that improvisation and playing-by-ear occur rarely (e.g., McPherson, 1993). Lessons tend to be dominated by teacher talk (Kotska, 1984; Persson, 1993; Sang, 1987; Tait, 1992), although only a small part of teachers' instruction is devoted to feedback (e.g., Goolsby, 1997; Speer, 1994; Yarbourgh & Price, 1989). The relatively few empirical studies available have suggested that instrumental teaching focuses on technique rather than on expression (see Hepler, 1986; Persson, 1993; Tait, 1992; Young, Burwell, & Pickup, 2003), and many method books for instrument teaching do not cover expressive aspects. In a study by Rostvall and West (2001), teachers did not consider expressive aspects at all.

Why is expression neglected in music education? As previously stated in the introduction, much knowledge concerning expression is tacit (Hoffren, 1964), and is therefore difficult for the teacher to convey in words to the student. There are also several myths about expression that might have had a generally negative impact on the teaching of expression in music education; for instance that (a) expression is a completely subjective entity that cannot be studied objectively, (b) you must feel the emotion in order to convey it to your listeners, (c) an explicit understanding is not beneficial to learning expression, (d) emotions expressed in music are different from everyday emotions, and (e) expressive skills cannot be learned (for a discussion, see Juslin, Friberg, Schoonderwaldt, & Karlsson, 2004).

These myths have several consequences for the teaching of expression. For example, consider myth (c) that partly stems from a general agreement that musicians are usually not aware of the details of how their expressive intentions are realized in their performance. This presents a problem for the teaching of expression, to the extent that it relies on verbal instruction (Tait, 1992). In addition, because expert performers do not consciously think about how to apply expressive features in their performance, one might wrongly conclude that students do not benefit from thinking about how to apply such features, even in the early learning stages.

If expression is completely subjective and has nothing to do with an explicit understanding or emotions as they are known, it will obviously be difficult to teach expression to students. A consequence of this view is that musical expression is wrongly believed to reflect only musical talent and thus cannot be learnt (Sloboda, 1996). These myths, however, do not imply that teachers never address expression in their teaching. When teachers do consider expression, the most common strategies are metaphors, modeling, focus on felt emotion, and verbal instruction.

## Traditional teaching strategies

*Metaphors*. A common strategy aimed at teaching expressive skills, is metaphors. Metaphors are used to focus the emotional qualities of the performance by serving as a reference or evoking a mood within the performer (Barten, 1998). However, although this strategy can be effective, metaphors depend on the performer's personal experience and images, and because different performers have different experiences, metaphors are frequently ambiguous (Persson, 1996).

*Modeling*. The teacher's performance provides a musical model of what is desired from the student, and the student is required to learn by imitating this model (Dickey, 1992). Modeling may be useful (Ebie, 2004), but it could be hard for the student to know what to listen for, how to represent it in terms of specific skills, and how to apply the skills in new situations. A pianist cited in Tait (1992, p. 528) says: "I try to avoid playing for students. […] Once they've copied the sound, they just repeat it and it doesn't mean anything."

*Focus on felt emotions*. This strategy means to focus on the performer's felt emotions, trusting that these emotions will automatically translate into the appropriate sound properties (Woody, 2000). However, felt emotion does not guarantee that the emotion will be communicated to listeners – neither is it necessary to feel an emotion in order to communicate it successfully. A problem noted by Sloboda (1996) is that students rarely monitor the expressive outcome of their own performances. Instead, they monitor their own intention and "take the intention for the deed" (p. 121).

*Verbal instruction*. Another common strategy is verbal instruction; that is, comments that directly address the relevant acoustic features (e.g., Woody, 1999). In order for this strategy to be successful, it requires that the teacher has explicit knowledge about expression, which may not always be the case.

## The importance of informative feedback

What is required for effective learning to occur? Based on an extensive overview of research on skill acquisition, Ericsson, Krampe, and Tesch-Römer (1993) proposed three elements that are required in a learning task for it to qualify as deliberate practice: (a) a well-defined task, (b) informative feedback, and (c) opportunities for repetition and correction of errors. They further claim that without adequate feedback, efficient learning is impossible and improvement only minimal, even if the motivation to learn is high.

Feedback has been formally defined as:

> "a process by which an environment returns to individuals a portion of the information in their response output necessary to compare their present strategy with a representation of an ideal strategy" (Balzer, Doherty, & O'Connor, 1989, p. 412).

This definition suggests that many traditional teaching strategies (e.g., metaphors and felt emotion) do not provide informative feedback, because they do not allow the performer to make a direct comparison of his or her current performance strategy with an optimal strategy. In her practical guide to better teaching and learning in instrumental teaching, Hallam (1998) argues that just telling the student that the performance is good or bad will not help him or her to improve – the feedback should be detailed. Tait (1992) concludes in his review that "teaching strategies need to become more specific in terms of tasks and feedback" (p. 532).

To summarize, expression is commonly neglected in music education, for various reasons mentioned above. One possible explanation for the deficits in music education may be that research on expression has not been able to provide music teachers with any theories that may guide the development of strategies for teaching expression.

# Expression in earlier research

## What is expression?

To begin with, it should be noted that the word 'expression' has been used in several different ways in the literature. In studies of music performance, 'expression' has been used to refer to the systematic variations in timing, dynamics, timbre, and pitch (i.e., the microstructure of a performance) that differentiate it from other performances of the same music (Palmer, 1997). 'Expression' has also been used to refer to the emotional qualities of music as perceived by listeners (Davies, 1994), or to a musician's sensitivity and skill to play a certain phrase in the exact right way (e.g., 'playing with great expression', London, 2002). Hence, one problem in previous research on expression is that there is a lack of definitions of expression. For instance, it is unclear how the different uses of the word relate to each other.

In addition to the different uses of the word, Juslin (2003) argues that a major problem in earlier research is a common tendency to regard expression as a single entity, a homogeneous natural entity. He argues that many studies have treated expression as a mysterious quality, of which there is simply 'more' or 'less', without specifying what is meant by the term. There is no consideration of *what* is expressed, or *how* it is expressive, which implies that there is only one way of performing expressively – by 'appropriate expressive deviations' (Juslin, 2003).

## How has expression been studied?

### Research on emotional expression

According to Gabrielsson and Juslin (2003), empirical research on emotional expression in music started already in the late 19th century, but became more frequent in the 1930s. The purpose was to investigate listener agreement on perceived emotional expression, or to investigate what factors in the musical structure according to the musical notation influenced perceived expression. Participants listened to pieces of music and reported perceived emotional expression by means of (1) free descriptions, (2) choice among descriptive terms, or (3) ratings of how well such descriptive terms applied to the music in question. This research has revealed a lot about how aspects of musical compositions affect listeners' emotional responses to music, but less about how different *performances* influence listeners' responses. Still, the same notated structure can be performed in several different ways, and the precise way it is performed may influence the listeners' impression of the music in profound ways.

**Research on music performance**

In previous research on music performance, psychologists have tended to approach performance expression mainly by measuring various acoustic variables of music performances (Gabrielsson, 1999, 2003; Palmer, 1997). Often a purely descriptive approach has been taken, in accordance with the tradition first established by Carl Seashore (1938). Back then, the data were presented in so-called 'performance scores', and expression was regarded as a 'deviation from the exact' (Gabrielsson, 1999). However, Seashore did not present any theory or provide any explanation of the relationships between deviation and affective response, which meant that the notion of expression was poorly conceptualized. Expression was reduced to tables or graphs of acoustic data, and the question of what these data in tell us about the origins of musical expression was somehow lost (Juslin, 2003). Some of the recent research based on neural network models, in which the computer discovers significant regularities in empirical data through inductive machine learning (Widmer & Goebl, 2004), is in a similar vein. It reflects Seashore's approach in that it involves no psychological theory and does not relate the description of the performance to the performer's intentions or the listener's perception.

In summary, earlier studies of music performance have focused on performance structure and ignored emotion, whereas earlier studies of emotion have focused on compositions and ignored performance aspects. Another problem is the tendency to regard expression as a single entity. There is thus a need for a definition of expression and for an integrative model that takes all aspects of expression into account. In addition, earlier research on expression in music performance has usually studied expression in isolation from any concerns with teaching of expression (Palmer, 1997), whereas studies of teaching strategies have not investigated the nature of expression itself (Marchand, 1975).

# A novel empirically-based approach

This section presents a novel and empirically-based approach to expression in music performance that could help to provide a solid foundation for the teaching of expressive skills in music education.

## Working definitions

Expression, communication, and emotion are complex concepts that are not easily defined. However, working definitions of how each term is used in the present thesis are provided in the following:

*Expression* refers to a set of perceptual qualities (e.g. structural, emotional, motional) which reflect psychophysical relationships between 'objective' properties of the music, and 'subjective' (or objective but partly person-dependent) impressions of the listener. Expression does not reside solely in the acoustic properties of the music (different listeners *may* perceive the expression differently), nor does it reside solely in the mind of the listener (different listeners *usually* agree about the general nature of the expression in a music performance). Expression depends on both these factors, in ways that can be modeled in a systematic fashion (Juslin, 2003).

*Communication* (of emotion, for instance) includes a performer's intention to *express* a specific concept and *recognition* of this concept by a listener. The performer may wish to emphasize an emotional character that is latent in the composition. The extent to which performer and listener agree about the emotional expression of the performance could pragmatically be seen as a measure of the accuracy of communication (Juslin, 2001).

*Emotion* is generally believed to consist of several components; cognitive appraisal, subjective feeling, physiological arousal, emotional expression, action tendency, and regulation. The focus in the present thesis is on the *expression* of emotion. The expression of an emotion can be spontaneous, as in 'I feel happy!', or the expression can be of a more symbolic or cognitive nature, such as when an actor or a musician portrays an emotion. There is also a difference between *emotion perception* (to recognize an emotion) and *emotion induction* (to be emotionally affected) that should be acknowledged (Gabrielsson, 2002; Juslin & Sloboda, 2001). The focus in the present thesis is on emotion expression (in the symbolic sense) and emotion perception, as two parts of a communication system.

## The GERMS model

In the past, the expression area has been investigated using a number of different approaches, as explained earlier. At the same time, expression has been viewed as a single entity and equated with everything that might be good in a performance (the conception of good not being specified). Juslin (2003) argues that psychology can make a contribution to how expression is conceived by considering the causes of the behavior in the performer and how to make sense of this behavior from what we know about humans. From a careful review of the literature, he suggests that performance expression is better thought of as a multi-dimensional phenomenon, consisting of five components of expression that are collectively referred to as the GERMS model. More specifically, the microstructure of a performance consists of the following five components:

1. *Generative rules (G)* that function to clarify the musical structure (Clarke, 1988). By means of variations in timing, dynamics, articulation, a performer is able to communicate group boundaries (Gabrielsson, 1987), metrical accents (Sloboda, 1983), and harmonic structure (Palmer, 1996).
2. *Emotional expression (E)* that serves to convey intended emotions to listeners (Juslin, 1997). By manipulating overall features of the performance such as tempo, timbre, and loudness, a performer is able to play the same structure with different emotional expressions.
3. *Random variations (R)* that reflect human limitations with regard to motor precision (Gilden, 2001). Even experts have trouble playing perfectly even time intervals (e.g., in a tapping task), there are still small and involuntary fluctuations in timing in their performance.
4. *Motion principles (M)* that prescribe that certain aspects of the performance (e.g., tempo) should be shaped in accordance with patterns of human movement (e.g., gesture). Shove and Repp (1995) argued that a pleasing performance is one in which the microstructure is shaped in accordance with basic constraints of animate motion (biological motion).
5. *Stylistic unexpectedness (S)* that refers to performer's deliberate attempt to deviate from stylistic expectations regarding performance conventions in order to add tension and unpredictability to the performance (Meyer, 1956).

In short, according to the GERMS model a performance should (1) convey the structure of the music, (2) express emotions, (3) exhibit motor precision, (4) be suggestive of human motion and gesture, and (5) deviate from stylistic expectations in aesthetically pleasing ways. All of these components reflect psychophysical relationships between acoustic features of the performance (specific patterns of information) and psychological characteristics of the listener (e.g., categorical perception). All five components occur together in complex interactions, but for certain purposes, such as research or teaching, it may be useful to treat them separately, because they have different origins, display different characteristics, and are processed partly by different regions of the brain (Juslin, 2003).

In a study by Juslin, Friberg, and Bresin (2002), the first four components (GERM) of the model were implemented and tested (the component S was not included because an implementation of this did not exist at the time). Synthesized expressive performances of a piece of music created by rules featured in the GERM model were rated by listeners. The results showed that each component contributed to the rated expressivity of a performance, and that they had predicted effects on listeners' impressions of the performance. The component that produced the largest effects on listeners' judgments of expressivity was E (Emotional expression).

These results are consistent with questionnaire data, which showed that 135 music conservatoire students defined expression primarily in terms of communication of emotions, as indicated by their free responses (Lindström, et al, 2003); 83% of the musicians claimed that they try to express specific emotions in their performance 'always' or 'often'. Furthermore, in a survey study by Minassian, Gayford, and Sloboda (2003), 53 high-level classical performers defined 'optimal' performances as ones where the performer (a) had a clear intention to communicate, usually an emotional message, (b) was emotionally engaged in the music, and (c) believed the message had been received by the audience. It seems safe to conclude that communication of emotions is a crucial aspect of music performance that a musician should address in order to be successful. Thus, the focus in the present thesis is on the emotional expression component.

# Emotional expression

Research on emotional expression, as manifested in performance features (e.g., timbre, timing) as opposed to features of the musical composition (e.g., harmony, melody) has matured in the last decade. Before the mid-1970s, there was not a single study of emotional expression in music performance. In 2001, there were about 30 studies of emotional expression in performance (Juslin, 2001). These studies covered a broad range of musical styles such as classical music, opera, folk music, jazz, and pop/rock. The most commonly studied emotions were happiness, sadness, anger, fear, and tenderness. The instruments represented in the studies were the violin, flute, clarinet, electric guitar, piano, trumpet, drums, synthesizer, and singing voice.

Most studies used a procedure in which music performers were asked to play a number of brief melodies to express different emotions (e.g., sadness) chosen by the researcher. These performances were recorded and used in listening experiments in order to check whether listeners could accurately recognize the intended expression. Each performance was also analyzed to study what means the performer used to accomplish each expression. It was assumed that, because the melody remained the same in different versions, whatever effects that appeared in listeners' ratings or in acoustic measures should mainly be the result of the performer's expressive intention.

The primary question for researchers was whether performers are able to communicate emotions to listeners at all. Kotlyar and Morozov (1976) first showed that listeners could successfully recognize opera singers' intentions of joy, anger, sorrow, and fear. Since then, several studies have shown that professional performers are able to communicate basic emotions (happiness, sadness, anger, fear, tenderness) to listeners with an accuracy as high as in facial and vocal expression of emotions (Juslin & Laukka, 2001, 2003).

If the purpose is to study communication as a process, it is also necessary to consider its mechanisms – particularly the *code* (i.e., the acoustic means) that performers use to convey each emotion. Several studies have indicated that performers manipulate acoustic variables such as tempo, sound level, articulation, timbre, timing, attack, decay, intonation, vibrato, and pauses to express different emotions (Juslin, 2001). Figure 1 shows how musicians use certain *cues* (pieces of acoustic information) in their way of expressing basic emotions. For example, expression of sadness is associated with slow tempo, low sound level, legato articulation, small articulation variability, slow tone attacks, and soft timbre ('soft' refers to little high-frequency energy in the spectrum), whereas expression of anger is associated with fast tempo, high sound level, staccato articulation, large articulation variability, fast tone attacks, and sharp timbre ('sharp' refers to much high-frequency energy in the spectrum). The expression created by overall features such as tempo can be further enhanced by accents on specific notes in the melodic structure (Lindström, 2003). The number of cues available depends on the musical instrument used. According to Juslin and Laukka (2003), the patterns of acoustic cues used to convey each emotion derive mainly from the nonverbal features of emotional speech.

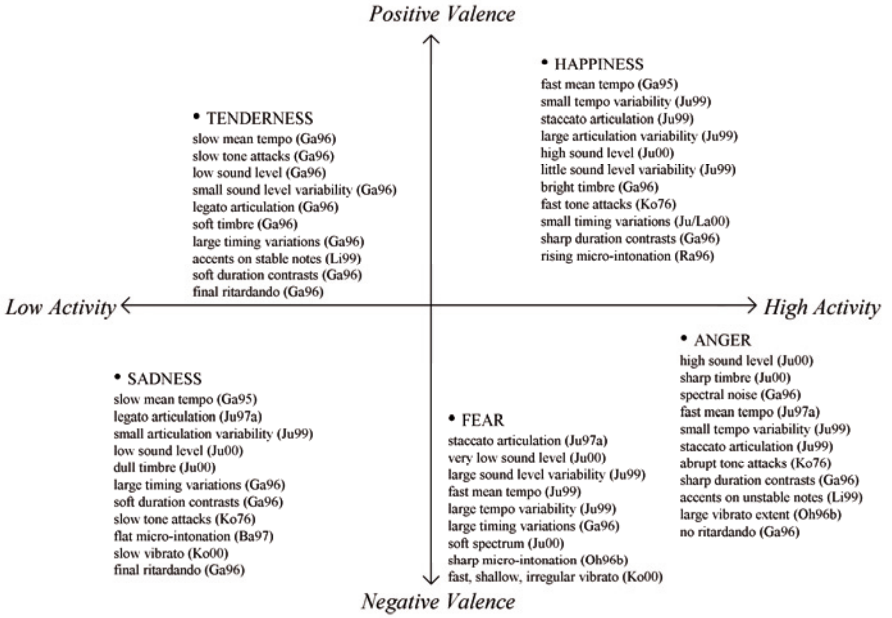


*Figure 1.* Summary of cue utilization in performers' communication of emotions in music. One representative study is cited for each cue. Authors' name are abbreviated to the initial two letters of the first author and the second author if needed for clarity, and the publication year to the last two digits, e.g., Ko76 = Kotlyar & Morozov 1976 (adapted from Juslin, 2001).

A peculiar phenomenon is that the communication of emotions in music performance is generally successful despite individual differences in the use of acoustic cues among both performer and listeners, and despite the fact that different instruments provide different cues at a performer's disposal. Juslin (2000) suggested that these aspects can be explained in terms of a modified version of Brunswik's (1956) lens model.

## The lens model

Communication of emotions means that there is both a performer's *intention* to express an emotion, and *recognition* of this same emotion by a listener. The performer might want to bring out joy in a certain passage of a piece of the music. One way of capturing the characteristics of the communicative process is presented by Juslin (2000, 2001) in terms of a variant of the *lens model*, originally presented by Egon Brunswik (1956) as a model for visual perception. The lens model (see Figure 2) illustrates how performers express (encode) emotions by means of a set of cues (e.g., variations in tempo, sound level, timbre) that are *probabilistic* (i.e., uncertain) and partly redundant. The emotions are recognized (decoded) by listeners who use these same cues to infer the expression.

The cues are probabilistic in that they are not perfectly reliable indicators of the intended expression (e.g., a fast tempo is not perfectly correlated with expressions of happiness, because fast tempo is also used in expressions of anger). No cue is completely reliable in isolation. Performers and listeners need to combine the cues in order for reliable communication to occur. In this regard, the communication process is not just simply a matter of pattern matching. The cues contribute in an additive fashion to listeners' judgments – each cue is neither necessary nor sufficient, but the larger the number of cues, the more reliable the communication. The redundancy among the cues partly reflects how sounds are produced on instruments. Thus, for example, a harder string attack produces a tone that is both louder and shaper in timbre.

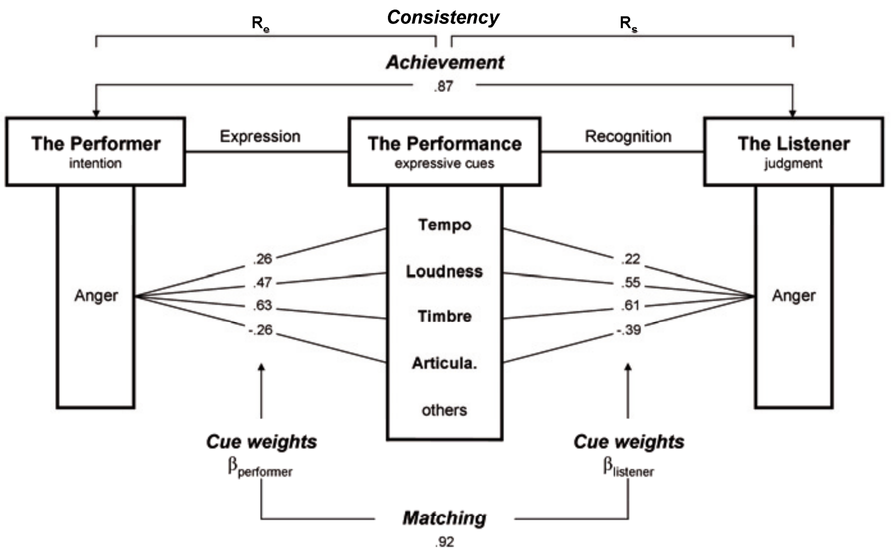


*Figure 2.* A Brunswikian lens model of emotional communication in music performance. The cue weights should be interpreted as follows: positive (as opposed to negative) signs indicate, respectively for each cue, fast (versus slow) mean tempo, high (versus low) sound level, sharp (versus soft) timbre, and legato (versus staccato) articulation (adapted from Juslin, 2000).

The key indices in the lens model describing the communicative process are:

- *Achievement* ($r_a$) refers to the relationship between the performer's expressive intention (e.g., intending to express anger) and the listener's judgment (e.g., perceiving anger). It is a measure of how well the performer succeeds in communicating a given emotion to listeners.
- C*ue weight* ($\beta_1$, $\beta_2$, $\beta_3$ …) refers to the strength of the relationship between an individual cue (e.g., tempo), on the one hand, and a performer's intentions or listeners' judgments on the other. Cue weights indicate how the individual cues are actually used by performers and listeners, respectively (e.g., that the performer uses fast tempo to express anger or that listeners use fast tempo to recognize anger).
- *Matching* (G) refers to the degree of similarity between a performer's and a listener's use of cues, respectively. For effective communication to occur, the performer's use of cues (i.e., his or her cue weights) must be reasonably matched to the listeners' use of cues.
- *Consistency* ($R_e$ and $R_s$) refers to the degree of consistency with which the performer and listeners, respectively, are able to use the cues. Other things equal, the communication will be more effective if the cues are used consistently.

The relations among the different indices of the lens model have been mathematically formulated in terms of *the lens model equation* which allows one to explain achievement in terms of matching and consistency (see Juslin, 2000, for a more elaborate explanation). The upper limit of communication accuracy is set by the degrees of matching, performer consistency, and listener consistency. If the emotional communication is not successful, this may be because (1) performer and listeners use the cues differently (poor matching), (2) the performer uses the cues inconsistently, and (3) the listener uses the cues inconsistently. By analyzing these indices separately, it is possible to interpret the success of communication in a particular situation. This information is needed in order to be able to improve the communicative process (Juslin, 2000).

The lens model implies that perfect accuracy of communication cannot be expected (if cues are uncertain, the communication process is uncertain too); the extent to which the expression of a piece of music is recognized depends equally on the sender and the receiver; and more than one cue utilization could lead to the same level of accuracy (because cues are partly redundant). Because there is no pressure toward uniformity in cue utilization, performers can communicate successfully with listeners without having to compromise their unique playing styles.

## Criteria for a useful teaching strategy

To achieve a creative interpretation of a piece of music, the performer may wish the emotional expression to be clear or ambiguous, stable or variable, specific or general. Only detailed knowledge about the relationships among expressive feature and their perceptual effects will help the performer to achieve the desired effects on listeners reliably. A useful teaching strategy aimed at emotional expression should improve emotional communication effectively and provide performers with the tools they need to develop their own personal expression. Given these aims, the teaching strategy should:

- Be well suited to the nature of the communicative process as described in research. A useful teaching strategy should take the acoustic cues that are used tacitly by performers and listeners into account and help render the communicative process transparent.
- Include the three elements for deliberate practice (Ericsson et al., 1993): (1) a well-defined task, (2) informative feedback, and (3) opportunities for repetition and correction of errors.
- Allow the performer to compare his or her playing with an 'optimal' or reference model, in accordance with the definition of feedback earlier.
- Relate sound properties of the performance to experiential concepts (e.g., emotion) relevant in interpretation. Modeling focuses on acoustic aspects, whereas metaphor focuses on experiential aspects; a useful strategy should resolve this dualism by describing the relationships between the two.
- Have its efficacy empirically demonstrated.

A teaching strategy that meets these criteria is presented below.

## Cognitive feedback

The lens model offers a useful tool for improving the communication of emotions in music. As described earlier, emotional communication involves a number of acoustic cues used by performers and listeners. Both expression and recognition of emotion are based on these cues. The lens model provides explicit knowledge about the *relationships* among performers, cues, and listeners. This information can be used to provide *cognitive feedback* (CFB), in which the performer can compare his or her playing to an 'optimal' way of playing (Juslin, 1998).

The term CFB was introduced by Hammond (1971) in studies of human judgment that provided judges with feedback about task properties and judgment strategies. CFB is usually contrasted with the notion of *outcome* *feedback*, which means that judges only receive information about whether the judgment was good or bad, but no information about why.

CFB differs from the kind of feedback commonly provided by music teachers in that it provides a direct comparison of the present strategy with an optimal strategy. Also, CFB differs from teachers' feedback in how the feedback is derived. Performers' manipulations of acoustic cues could be audible to a listener, but it is difficult for a human perceiver to infer the statistical relationships that exist among expressive intentions, acoustic cues, and listener judgments. It is often the case that judges are unable to explain the basis of their judgments – especially in situations that feature several uncertain cues (Brehmer, 1994). However, CFB takes care of this problem by using a statistical method (i.e., multiple regression analysis) that makes it possible to explicitly describe the complex relationships with a precision that would be hard to achieve for a human perceiver.

In a study by Juslin and Laukka (2000), guitar players increased their communication accuracy by 50% after only one CFB session, as indicated by listening tests. In the study, regression models of the performers and the listeners were obtained by means of manually analyzing the acoustic cues of the performances and conducting regression analyses. Such measurements and analyses are very complex and time-consuming, and are not really an alternative for teachers to use in music education. However, CFB has now been implemented in computer software – the Feel-ME program.

## The Feel-ME program

The Feel-ME program is an application that can automatically analyze music performances and offer CFB to performers. A first prototype of the program was implemented in *Matlab*©. The purpose was to create a program that would be easy to use even for music students with little experience of using computers. The basic procedure is that the performer records performances with various emotional expressions, and then is presented with CFB that contains suggestions on how to change certain acoustical aspects in order to enhance the expression. The performer then again records performances according to the feedback suggestions, and by retrieving CFB again he or she is able to see whether the communication accuracy improved or not.

In the first phase, the performer is instructed to record a number of different performances of the same melody in order to communicate various emotions (e.g., happiness, sadness, anger) that are selected at the start. The performer records several versions of each emotional expression to obtain a representative sample of performances. The performances are stored in the computer memory, and acoustic cues (e.g., tempo, sound level, articulation) are automatically analyzed by the program. The recording is first segmented into tone boundaries from an analysis of sound level and pitch. Potential tone onsets and offsets are detected by identifying segments of similar pitch and substantial dips in the sound level. For each detected tone, the following cues are computed: pitch (in semitones), sound level (dB, upper quartile of sound level within onset-offset), instantaneous tempo (notes per second), articulation (percentage of pause duration), attack velocity (dB/s), spectral content (dB, difference between high and low spectral content, a correlate of perceived timbre), vibrato rate (Hz), and extent of vibrato (semitones). Then, multiple regression analysis is used to model the relationships between the performer's expressive intentions and the acoustic cues. This produces indices of consistency and cue weights for the performer. The performer's model is also compared to a stored regression model of listeners' judgments of emotion in music performances. This listener model was obtained from previous listening tests in which listeners were asked to judge the emotional expressions of a wide range of music performances with varying emotional expression. Statistical analysis was used to model the relationships between listeners' judgments and acoustic cues, thus producing a general model that may be used to simulate new listener judgments by statistical bootstrapping. From this analysis, the program obtains indices of achievement, matching, consistency, and cue weights. This information is finally transformed to a graphical interface, according to set criteria (for further details, see Study II).

In the second phase, the performer requests feedback from the program; this includes a visual and numerical description of the performer's use of cues, the listeners' use of cues, the matching between performer's and listeners' cue weights, the consistency of the performer's use of cues, and the achievement. This is shown in a graphical interface that resembles the lens model, which makes it possible to directly compare how performer and listeners use the same cues in the performances (Figure 3). In order for the performers to easily understand their accuracy of communicating the emotions, the values showing achievement, matching, and consistency are transformed from correlations to scores from 1 to 5, based on the previous Swedish school system. The arrows indicate poor matching that signals that a change in the utilization of a cue in a specific direction is recommended. A recommendation is also expressed verbally (e.g., slower). If the performer is using cues inconsistently when expressing an emotion, this will show in the consistency index. From this point, the performer should try to change his or her use of the cues according to the provided feedback (e.g., to play slower to communicate sadness).

In the final phase, the performer repeats the task once again (i.e., records a number of performances that express the specific emotions). The program again records and analyzes the performances, and uses simulated listener judgments to provide updated lens model indices. Results from successive feedback cycles are stored in the program and can be retrieved at will from a database to plot learning curves for achievement, matching, and consistency, as well as changes in cue weights. The goal is to see whether the performer improves his or her communication by changing the use of cues in the ways recommended by the CFB. When observing the updated CFB, the performer can easily see which cues are used effectively and which cues need continued attention. This feedback cycle might be repeated as many times as desired. Although there exist a large number of computer programs for the music profession (for overviews, see Bartle, 1987; Webster, 2002), this is the first program aimed specifically at enhancing emotional expression.

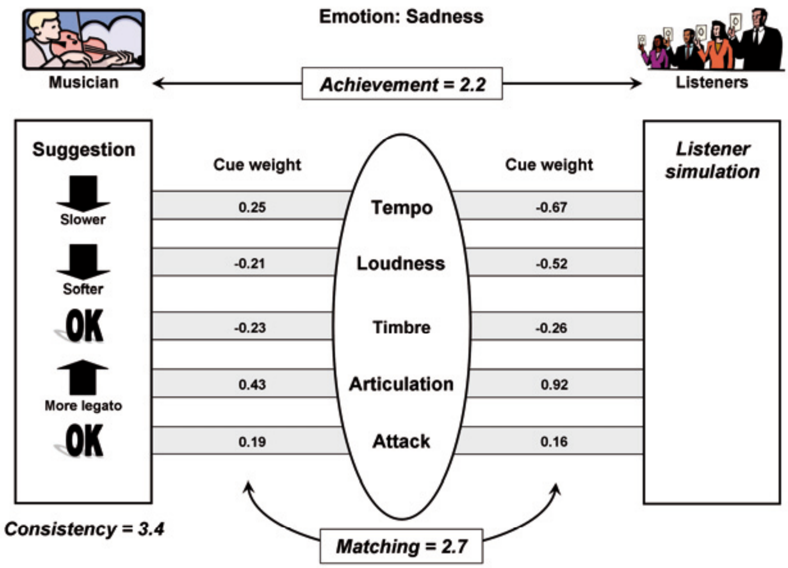


*Figure 3.* The graphical interface for CFB in the Feel-ME program (adapted from Juslin et al., 2004).

# Aims of the thesis

The overall aim of this thesis was to develop and evaluate a new method for teaching emotional expression in music performance based on psychological theory and research. This task was carried out in three empirical studies with their own specific aims:

Study I is an observational study of music teaching at different levels of education. The goal was to investigate the nature of instrumental teaching in a natural setting with a special focus on expression, trying to pinpoint what is wrong with current teaching.

Study II presents a new computer program aiming to improve performers' emotional expression. The purpose of the study was to evaluate the program with regard to both its performance (does the program improve a performer's emotional expression?) and its usability (is the program user-friendly?).

Study III is a deception experiment, which aimed to investigate whether performers' views towards computer-assisted teaching of expression reflect general attitudes towards computers or preferences concerning the precise feedback contents.

# Study I

## Background and aims

A natural starting point for an attempt to improve teaching of expression is to obtain more information about how expression is currently being taught. This is an important task, because the best solution to the problem depends on what exactly is wrong with the current teaching. However, real-world investigations of instrumental teaching are still rare, and what goes on in individual lessons is largely hidden from view (Hallam, 1998). Not only does instrumental teaching in its natural context involve complex social situations whose meanings are very difficult to capture using traditional experimental methods, teachers also avoid participating in field studies because there is an element of risk in exposing their behavior to an observer. In his review of teaching strategies and styles, Tait (1992) argues that there is a need for detailed descriptive research to get a clearer understanding of music teaching.

In order to get an understanding of the nature of instrumental teaching, it needs to be studied in its natural context. However, real-world interactions contain too much information to be reduced to a few quantitative measures. It can be argued that an understanding of the complex social interactions in instrumental teaching requires micro-analyses of teaching sessions, using both qualitative and quantitative methods. Qualitative methods emphasize the qualities of processes and meanings in natural settings (e.g., Denzin & Lincoln, 2000), whereas quantitative methods may be useful to effectively describe structural aspects of the teaching process.

In Study I, the aim was to explore the nature of instrumental music teaching in its natural context, with a focus on expression and emotion. Specifically, the study aimed to provide information regarding the following questions:

- What is the nature of the social interaction during teaching?
- How much time is allocated to various teaching activities?
- What features of performance does the teaching focus on?
- Are issues of expression and emotion explicitly addressed?
- What types of feedback strategies are used by the teachers?
- To what extent is teaching guided by explicit goals or plans?

# Method

## Participants

The study consisted of a series of case studies, with each music teacher representing a case and his or her students representing 'cases within cases'. Five instrumental teachers, aged 30-70 (mean age = 50.4 years), four males and one female, participated in the study. They were teaching the viola, the guitar, or singing at different educational levels in Sweden, and their teaching experience ranged from 3 to 40 years (mean teaching experience = 25.4 years). Twelve music students, aged 15-35 (mean age = 23.5 years), six females and six males, took part in the study. Their playing experience ranged from 3 to 22 years (mean playing experience = 9.25 years). Eight of the participants studied at conservatory level, two at upper secondary level, and two at a municipal music school.

## Procedure

Video observation was used in order to study instrumental teaching in the field. Music teachers and students were thus filmed during lessons at their school using a digital video camera. Participants were informed that the recording would be used to investigate the nature of instrumental teaching, that it would be used for research purposes only, and also that they could abort the recording at any time. After the lesson, a questionnaire on various background variables including teachers' and students' views on expression was administered.

## Analyses

Instrumental lessons were transcribed, content-analyzed, and coded into categories of feedback and language use. First, a draft transcription of each lesson was carried out (e.g., teacher talk, student talk, student performance, teacher performance, and other behaviors including gestures), then a content analysis of the conversation was conducted. The transcription was coded by two independent coders, using a pre-determined codebook for language use from a study by Rostvall and West (2001, pp. 81-82). The codebook was developed for the purpose of analyzing the communicative functions of language during instrumental teaching, and was largely based on Rostvall and West's own experience as instrumental teachers. Inter-coder agreement was estimated using Cohen's Kappa (e.g., Howell, 1992, p. 148), and mean inter-coder agreement across cases and coding categories was $\kappa = .90$.

Quantitative data were used to describe structural characteristics of the instrumental lessons (e.g., time allocation to various activities, language use, frequency of use of specific feedback strategies, word counts). Finally, a qualitative analysis of the complete transcriptions was carried out to discover particular teaching sequences and themes that could be used to characterize the interaction between the teacher and the student (Ryan & Bernard, 2000; Silverman, 2000).

# Results

## Language use and feedback strategies

Measures of time allocation to various activities indicated that lessons were dominated by talk (61%), rather than by playing (39%). Word counts based on the transcription of verbal behavior showed that the teacher accounted for 73% (range = 62-88%) and the student for 27% (range = 12-38%) of the words uttered during lessons.

Language use was coded in terms of five educational functions: *Testing* (e.g., questions such as 'should I continue from here?'); *Instructional* (e.g., instructions, evaluations, such as 'play it from the top', 'that was excellent - bravo!'); *Analytical* (longer sections of coherent reasoning in several steps, which serve to explain 'cause-and-effect' relationships, such as 'in order to play this phrase, you must use this type of fingering, because otherwise you will not be able to change string and play the following phrase in time'); *Accompanying* (utterances that mainly serve to guide the conversation by confirming or objecting to earlier utterances, such as 'OK', 'Yes, that's right'); and *Expressive* (utterances that explicitly address expressive aspects of the performance, such as 'more expression!').[1] The overall language use for teachers and students during lessons is shown in Figure 4.

As can be seen, the teachers' language use was mainly instructive and to a lesser extent accompanying (see Figure 4), whereas the students' language use shows the reverse pattern. There were rather few occurrences of testing, analytical, or expressive language use. These results suggest that the typical pattern in instrumental teaching consists of the teacher telling the student what to do and the student responding briefly to these instructions. The results further suggest that questions, analyses, and discussion with regard to expressive aspects of performance occur rarely.

[1] Our use of the coding category *expressive* differs somewhat from the use in Rostvall and West (2001) in that we focus on whether the utterance has as a function to say something about the expression of the performance, whereas Rostvall and West focused on whether the utterance itself was 'expressive' also (as in a metaphor). The former use of the category was regarded as more useful in the present study.

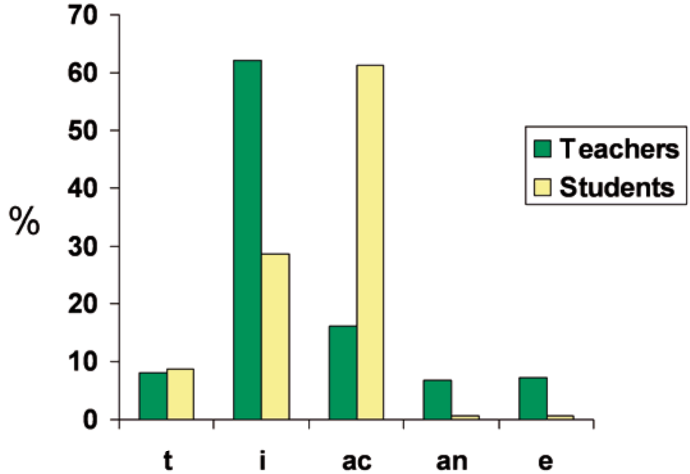


*Figure 4.* Overall language use (%) during instrumental lessons by teachers and students (t = testing, i = instructive, ac = accompanying, an = analytic, e = expressive)

The frequency with which different kinds of feedback were provided by the teacher was also investigated. Feedback was defined as sequences where the student plays something and the teacher responds to the playing in a way that allows the student to assess his or her progress. Feedback sequences were coded into four categories: (a) *verbal instruction* (the teacher offers comments on technical or acoustic features); (b) *modeling* (the teacher's performance provides a model of what is desired from the student and the student attempts to imitate the model); (c) *outcome feedback* (the teacher provides information about whether the performance was good or bad, albeit no information about why), and (d) *metaphors* (the teacher uses a simile or figure of speech that focuses the expressive qualities of the performance by serving as a reference). The results are displayed in Table 1.

The results indicate that the most frequently used feedback strategy was outcome feedback, followed by verbal instruction (see bottom of Table 1). Modeling and metaphors were less frequently used. There were individual differences among teachers, however: Although the metaphor strategy was least used by all, there were large differences regarding the use of outcome feedback and verbal instruction (e.g., for Teacher C, 16% of the feedback was verbal instruction and 58% was outcome feedback, whereas for Teacher E, 57% of the feedback was verbal instruction and 14% was outcome feedback).

Table 1. Results from the coding of teachers' feedback to students during lessons

| | Feedback (%) | | | |
|---|---|---|---|---|
| **Teacher** | **VE** | **MO** | **OU** | **ME** |
| **A** | 39 | 20 | 28 | 13 |
| **B** | 36 | 18 | 40 | 6 |
| **C** | 16 | 26 | 58 | 0 |
| **D** | 25 | 7 | 68 | 0 |
| **E** | 57 | 17 | 14 | 12 |
| **Mean (%)** | **35** | **17** | **42** | **6** |

*Note.* VE = verbal instruction, MO = modeling, OU = outcome feedback, ME = metaphor

## The teaching process

In addition to the overview of the basic structure of the analyzed lessons, an in-depth qualitative analysis was conducted by repeatedly going through the material to discover particular teaching sequences and themes that could be used to characterize the social interaction between the teacher and his or her students. Although there were individual differences among teachers and students, the overall teaching approach was remarkably similar:

- Lessons were mainly devoted to the reproduction of specific pieces of music from a written notation
- The focus was mostly on technique and on the musical score
- Playing-by-ear occurred rarely
- Improvisation did not occur at all
- Lessons were rarely guided by explicit goals or plans
- Lessons usually proceeded by asking the student to play practiced pieces while the teacher commented on the performance in an 'ad-hoc' manner

The general teaching approach was consistent with the 'master-apprentice' approach (Jørgensen, 2000) and was similar across all cases, which differed with regard to teacher, student, instrument, teaching level, and musical style.

Six themes occurred repeatedly during the lessons (presented in order of frequency): 'Technique' (focuses on technical problems of the performance, e.g., intonation, fingering), 'Notation' (focuses on some aspect of the music notation, e.g., different editions or analyses), 'Practical issues' (focuses on practical issues that concern the lessons, e.g., deciding about the next meeting), 'Expression' (focuses on expressive aspects of the performance, e.g., communication, emotions, interpretation, shaping), 'Self-confidence' (focuses on boosting a student's self-confidence through encouragement or challenge), and 'Instrument problems' (focuses on problems related to the student's instrument or ergonomics).

To summarize the results regarding expression, which was of special interest in this study, qualitative analyses of the material were consistent with quantitative data in suggesting that most of the relatively few teaching sequences that focused on expression involved implicit teaching strategies (such as modeling). When expression was explicitly addressed, teachers tended to use vague statements that provided little concrete advice to the student, such as: 'Play from the heart! Not the brain, it's the heart… sort of. But it's a combination'; 'You have to be moved yourself in order to move others'. Altogether, the results suggested that the teachers, rather than being guided by explicit or systematic knowledge about expression, attempted to convey an intuitive way of thinking about expression that served to help students discover expression on their own. (For more elaborate examples of individual teachers' teaching sequences and themes, see Study I.)

## Conclusions

The results from Study I indicated that instrumental teaching was mainly devoted to reproduction of particular pieces of music from a notation. The focus of the teaching was mostly on technique and on the written score, but expression was also addressed. However, expression was mostly dealt with implicitly rather than explicitly. Rather than to discuss specific emotional expressions, teachers used colorful but vague descriptions of expression that on some occasions led to misunderstandings. The most frequently used feedback strategies were outcome feedback and verbal instruction, rather than metaphors and modeling. The teaching was rarely guided by explicit goals, tasks, or systematic teaching patterns.

The findings illustrate positive as well as negative sides of 'traditional' instrumental teaching: on the positive side is the fact that teachers are able to consider several aspects of a performance simultaneously and can provide rich information to the student while the student is playing. On the negative side is the fact that teachers' feedback is not always provided in the most systematic manner, which means that much of the information may be lost during lessons.

As Duke (1999) points out, further research is needed concerning the efficacy of various teaching strategies regarding expression. As emphasized by Tait (1992), there is a need for the development of novel tasks and interventions that may enhance expression. The present results suggest that such performance interventions could benefit from the inclusion of explicit goals, systematic teaching patterns, and specific feedback.

# Study II

## Background and aims

As previously stated, although emotional expression is a crucial element in music performance, research has suggested that expressive skills are often neglected in music education. Study I further suggests that when expression *is* addressed by teachers, it is often done implicitly and without informative feedback. It has been suggested by several authors that learning of musical performance skills in general and expressive skills in particular benefits from clear instructions (Hallam, 1998; Woody, 1999), which suggests that implicit strategies are less effective.

Study II presents and evaluates a computer program that automatically analyzes music performances and provides specific feedback to musicians in order to enhance their communication of emotions. The program is based on a CFB approach, and provides feedback that includes suggestions on how to improve the communication. The purpose of the study was to evaluate the program with regard to both its performance (Does the program improve a performer's communication of emotions?) and its usability (Is the program user-friendly?).

Thirty-six guitar players were randomly assigned to one of 3 conditions: (1) feedback from the computer program, (2) feedback from music teachers, and (3) repetition without feedback. Based on the assumption that both the computer program and the music teachers would be able to provide useful feedback to the performers, it was predicted that groups 1 and 2 would show a greater improvement in communication accuracy than group 3. In addition, assuming that the program would provide more specific feedback than the music teachers, group 1 was predicted to show a larger improvement in communication accuracy than group 2. The predictions were tested in two listening experiments and also by performance measures obtained from the computer program.

Efficacy is not the only criterion in evaluating a new application, it is equally important to take the usability of the program into account. If people do not have favorable reactions to the application, they will not want to use it anyway. Video observation and a questionnaire were used to measure user interaction and subjective reactions to the program, respectively. Based on previous studies of musicians' attitudes towards computer-assisted teaching of expression (Lindström et al., 2003), we anticipated a negative impression of the program.

# Method

## Recording experiment

### Participants

There were 36 guitarists (mean age = 28 years), 35 males and 1 female, that participated in the study. Their playing experience varied from 5 to 39 years (mean playing experience = 16.5 years), and they mainly performed jazz and rock. Four guitar teachers (mean age = 38 years), all males, also participated in the study. Their teaching experience ranged from 6 to 30 years (mean teaching experience = 14.5 years) and all of them regarded it as important to teach expression to music students.

### Procedure

The guitarists were randomly assigned to one of three experimental groups:

- Cognitive feedback from the Feel-ME program (CFB group)
- Verbal feedback from a music teacher (Teacher group)
- Repetition without feedback (Contrast group)

The basic task was the same in all conditions. The performer was asked to play a brief melody, *When the Saints*, so that it would express *happiness*, *sadness*, *anger*, and *fear*, respectively. The performer was asked to play five versions of each emotion and to make them as similar as possible in order to get a reasonable number of cases for the program's statistical analysis. 20 performances were recorded in both a pretest and a posttest. Hence, each performer recorded 40 performances (i.e., 5 versions, 4 emotions, 2 tests). The recording process was handled by the experimenter, except for the CFB condition in which the performer interacted directly with the Feel-ME program. In order to avoid cognitive overload and ceiling effects, performers in the feedback groups where instructed to focus on two of the four emotions, namely those two emotions that the performer had been least successful in expressing initially (as revealed by the achievement score in the Feel-ME program), during the feedback session. The remaining features of the procedure were unique to each group, as explained below.

*Cognitive feedback group*. The performer first briefly explored the Feel-ME program, and was then required to go through one cycle of CFB. The feedback focused on four acoustic cues: tempo, sound level, articulation, and timbre. The performer's interaction with the program was videofilmed. After completing the CFB cycle, the performer filled out a usability questionnaire.

*Teacher feedback group*. The performer carried out the same basic task as in the CFB condition, except that the feedback was provided by a teacher. After the recording of the first 20 versions (the teacher was not present), the performer took a short break and left the laboratory. The teacher entered the laboratory, read the instructions, listened to the 10 target performances and wrote his feedback on a paper. The teachers could use any type of verbal instruction (metaphors, musical directions, focus of felt emotion, outcome feedback) to help the performer to enhance his or her communication of the two target emotions. Musical modeling (i.e., where the teacher plays an instrument) was not allowed. Finally the performer returned to the laboratory again and the teacher provided feedback to the performer as in a regular teaching session.

*Contrast group*. The performer received no feedback and performed the musical material twice (pre- and posttest), with a break in between. After the recording, the performer filled out a background questionnaire.

## Measures

### *Performance measures*

The Feel-ME program provided performance measures that were used to provide CFB to the performers. Acoustic measures of tempo, sound level, articulation, and timbre were automatically analyzed by means of the CUEX algorithm (see Friberg, Schoonderwaltd, & Juslin, 2007). The recording is segmented into tone boundaries through analyses of sound level and pitch, and for each detected tone, a number of parameters are computed. As explained earlier, statistical analyses were used to model the correlations between acoustic cues, the performer's expressive intentions, and listeners' judgments of the emotional expression. From this, the program obtained indices of achievement, matching, consistency, and cue weights (see p. 20).

### *Usability measures*

The user interaction was measured by means of video observation and a questionnaire. Regarding the observation, the performer's screen activity, speech, and behaviors such as facial expressions and postures were transcribed, and episodes of particular importance (e.g., mistakes) were finer transcribed. The questionnaire concerned such aspects as the naming of commands, the organization of the program modules, and more general impressions of the program.

## Listening experiments

### Listeners

Two listening experiments were conducted. Experiment 1 consisted of 16 musically trained listeners (mean age = 28 years), 9 females and 7 males participated. Experiment 2 consisted of 14 musically untrained listeners (mean age = 25 years), 7 females and 7 males participated.

### Music material

Altogether, the performers recorded 1440 performances (36 performers x 5 versions x 4 emotions x 2 tests) in the recording experiment. However, some reduction was necessary to obtain a manageable number of stimuli for a listening experiment. Therefore, we randomly selected one performance of each of the two emotions for each performer and test (36 performers x 2 emotions x 2 tests). This subsample of 144 performances was used in both listening experiments.

### Procedure

Both experiments were conducted individually, using computer software to play sound files (in randomized order) and to record listeners' responses.

In Experiment 1, listeners made forced-choice judgments of the 144 performances, which were presented in blocks of pairs with similar intended emotional expression. Unknown to the listener, one member of each pair was a pretest performance by one of the 36 performers and the other member was a posttest performance by the same performer. The listener's task was simply to judge which of the two versions in each stimulus pair was the most *happy*, *sad*, *angry*, and *fearful*, respectively.

In Experiment 2, the listeners were instructed to rate each stimulus with regard to how well it matched each of the adjectives *happy*, *sad*, *fearful*, and *angry* on a scale from 0 to 9 (0 = 'not at all', 9 = 'completely').

# Results

## Performance

As predicted, the results from Experiment 1 (Figure 5) suggested that the CFB and the Teacher group showed a larger improvement than the Contrast group. This effect was 'medium' ($r_{pb}$ = .34), according to Cohen's (1988) guidelines. The results also showed that the CFB group improved more than the Teacher group ($r_{pb}$ = .36).

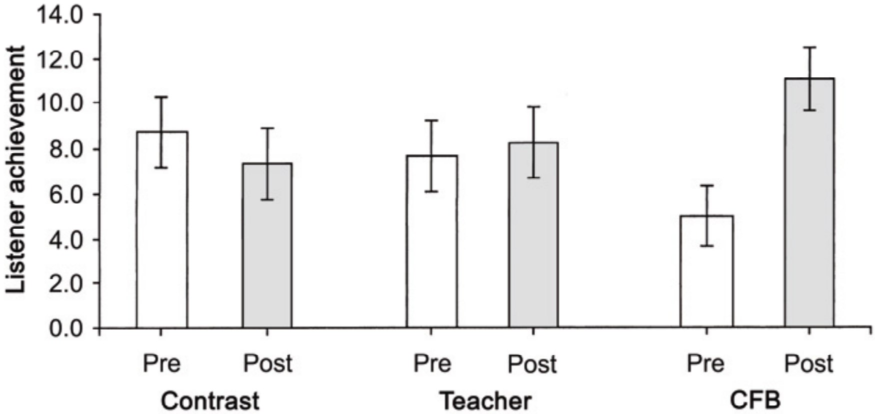


*Figure 5.* Listeners' forced choice judgments in Experiment 1 as a function of pre- (light bars) and posttest (dark bars) and experimental condition. Whiskers indicate 95% confidence intervals around the mean.

The findings of Experiment 1 were replicated in Experiment 2; see Figure 6. The effect of CFB was smaller though, perhaps because differences were more difficult to detect in the rating-scale task than in the forced-choice task.

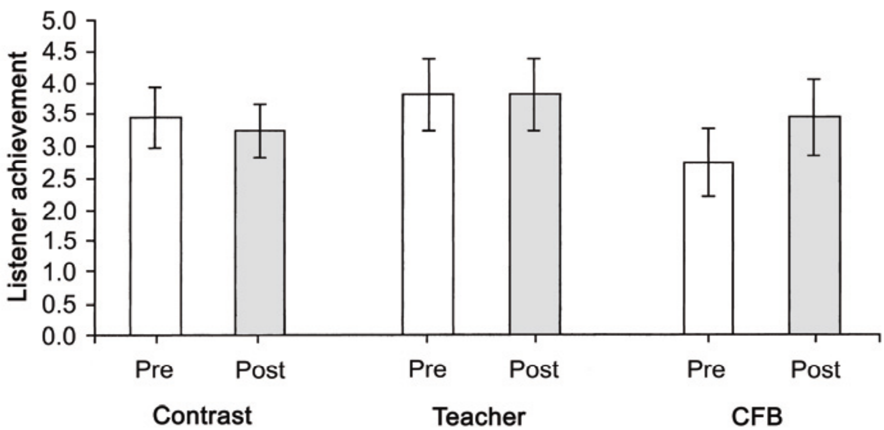


*Figure 6.* Listeners' mean ratings of intended emotions in Experiment 2 as a function of pre- (light bars) and posttest (dark bars) and experimental condition. Whiskers indicate 95% confidence intervals around the mean.

The results from the Feel-Me program itself was based on a larger sample ($N$ = 720) than the listening tests ($N$ = 144). The results are consistent with those of Experiments 1 and 2 in suggesting that both the CFB group and the

Teacher group improved more than the Contrast group, and that the CFB group improved more than the Teacher group (see Figure 7).

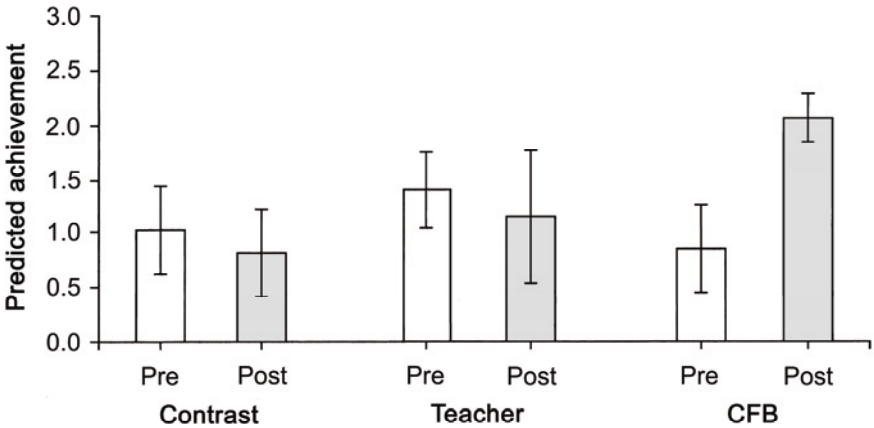


*Figure 7.* Predicted level of achievement by the Feel-ME program as a function of pre- (light bars) and posttest (dark bars) and experimental condition. Whiskers indicate 95% confidence intervals around the mean.

## Usability

The main results from the usability questionnaire indicated that the majority of the performers thought that the Feel-ME program was ‘rather good’ (75%). Furthermore, the performers rated the program as ‘very easy to understand’ (75%) and ‘very easy to learn to use’ (67%). Turning to the results from the video observation, results showed that the users found the overall design of the Feel-ME program very easy to understand. Most mistakes were related to a misinterpretation of the distinction between a *session* (one recording of a set of performances by a performer) and a *project* (a minimum of two linked sessions by the same performer).

A majority of the performers understood the feedback suggestions, but some of them (25%) reported that the feedback was difficult to understand (i.e., rating < 3 on a scale from 1 to 5). Also, 33% of the performers found it difficult to change their playing in accordance with the feedback suggestions – mainly because it was difficult to change one cue without unintentionally changing other cues.

Despite the mostly positive impressions reported, when asked whether they would consider using the program in the future, a majority (75%) of the performers responded negatively (i.e., rating < 3 on a scale from 1 to 5).

## Conclusions

Study II demonstrated that it is possible to create a computer program that automatically analyzes performances and provides musicians with feedback in order to enhance their expression of emotions. Both the Feel-ME program and the music teachers were more effective in enhancing the communicative process than simple repetition without feedback. The results also suggested that feedback from the program produced larger improvements in accuracy than verbal feedback from teachers.

Usability measures showed that the Feel-ME program was experienced positively by most of the users, but a striking finding was that most users responded negatively when asked whether they would consider using the program if they had the chance. Comments in the questionnaire suggested a negative attitude towards the use of computers to learn expression ('What does a computer know about emotions?'), while usability measures indicated the computer feedback itself could be perceived as difficult. Thus, Study II suggested the need to further investigate the reasons underlying performers' negative views on computer-assisted teaching of expression.

# Study III

## Background and aims

Study I suggested that expression is neglected in instrumental teaching, and that teachers rarely provide informative feedback to their students. In Study II, a new and empirically-based method aiming to improve performers' skills with regard to emotional expression was presented. The Feel-ME program efficiently improved performers' abilities to communicate specific emotions, but when asked, the performers did not want to use it in their own learning.

Could the negative impressions of the Feel-ME program in Study II be due to a generally negative attitude towards using computer with regard to expressive skills or due to a dislike of the actual feedback contents from the program? The purpose of Study III was to investigate this issue, an issue that has crucial implications for the development and application of new teaching methods concerning expression in general, and for the Feel-ME program in particular. More specifically, the aim was to investigate the effects of (a) the participants' beliefs about whether the feedback was produced by a teacher or a computer (feedback delivery), and (b) the feedback contents in terms of whether they were actually produced by a teacher or a computer (feedback production).

On the basis of previous findings that performers have negative attitudes towards computers in regard to expression (Lindström et al., 2003) and the results from the previous study that evaluated the Feel-ME program (Study II), we predicted that (1) performers would give higher quality ratings to the feedback if they believed that it originated from a music teacher (because of negative attitudes towards computers), and (2) performers would give higher quality ratings to the feedback if it was produced by a computer (because computer feedback is more focused and less vague).

To gain further insights about these quality ratings, the performers were also required to judge the comprehensibility and the level of detail of the feedback that they received. Furthermore, to obtain more efficient tests of the treatment effects, we measured a selection of background variables that might influence the performers' responses, such as musical experience, interest in computers, experience of using computers, attitudes towards the possibility to learn to play expressively. Selected background variables were included as covariates in the analyses.

# Method

## Participants

There were 80 guitarists (mean age = 25 years), 70 males and 10 females that participated in the study. Their playing experience varied from 2 to 50 years (mean playing experience = 11 years), and about half of them (56%) had some experience of formal tutoring on guitar.

There were also four guitar teachers (mean age = 29 years), all males, involved in the study. Their teaching experience ranged from 5 to 10 years (mean teaching experience = 6.75 years) and they were teaching guitar at different levels of education in Sweden.

## Procedure

The performers were randomly assigned to one of four experimental conditions. In a 2 x 2 between-subjects factorial design, we manipulated (a) the performers' belief about whether the feedback was produced by a teacher or a computer program (feedback delivery) and (b) the feedback contents in terms of whether they were actually produced by a teacher or a computer program (feedback production). The overall procedure was the same in all experimental conditions: the performers were asked to play a piece of music to express four emotions, received verbal feedback on their performances, and judged the quality of the feedback on rating scales.

The performer was required to play a brief melody, *Vem kan segla* (Finnish folk melody), so that it would express *happiness*, *sadness*, *anger*, and *fear*, respectively. The piece was chosen because it was short, familiar, and easy to play. The performer was asked to play three versions of each emotion (minimum requirement for the Feel-ME program to carry out performance analyses and provide feedback), and to make them as similar as possible. The recording process was handled by the experimenter and the performances were recorded by means of the Feel-ME program.

After the recordings, the performer took a break and left the laboratory. During the break and unknown to the performer, activities of producing feedback and preparations for feedback delivery took place. After some time, the performer returned to the laboratory and read instructions that acknowledged that he or she was going to be provided with feedback on his or her playing. The feedback instruction to the performer read: "You are now going to be provided with feedback on your performance from 'an experienced teacher' (Teacher-Teacher and Computer-Teacher condition) or 'an advanced computer program' (Computer-Computer and Teacher-Computer condition) with suggestions on how you can change your playing in order to enhance your communication of emotions to listeners". In the Teacher-Teacher and the Computer-Computer condition, the instructions provided to the performer were true. However, for the Computer-Teacher and the Teacher-Computer condition, the instructions were false.

After the feedback session, the performers rated the dependent variable *feedback quality* on a scale from 0 to 10 in a questionnaire. In order to obtain further explanation and insight into the main result of rated feedback quality, the performers were also required to rate *feedback comprehensibility* and *feedback detail level* on a scale from 0 to 10. They concluded the experiment by filling out a background questionnaire. Afterwards, the performers were instructed not to reveal the design of the experiment. The remaining features of the procedure were unique for each condition group:

*Teacher-Teacher condition*. The music teacher came to the laboratory and listened to the target performances in order to write down feedback to the performer. As in Study II, teachers were asked to use verbal instruction. When ready, the teacher temporarily left the laboratory for a few minutes while the performer returned from his or her break and read the feedback instructions. Then, the teacher entered the laboratory to read the suggestions aloud to the performer. The paper with the feedback was then handed over to the performer, and the teacher left.

*Teacher-Computer condition*. The music teacher carried out the same basic task as is the Teacher-Teacher condition, except for the following; the teacher produced feedback, left the laboratory and (unknown to the teacher) the experimenter transferred the teacher's written feedback to a PowerPoint document designed to look as a screen dump picture from a computer program. After the performer returned from his or her break and read the instructions, the experimenter read the feedback suggestions aloud to the performer, and then handed over the paper with the suggestions.

*Computer-Computer condition*. The procedure in this condition was the same as in the Teacher-Computer condition, except that there was no teacher involved (i.e., the feedback from the computer was written in the PowerPoint document mentioned above). As in Study II, the computer feedback focused on four acoustic cues: tempo, sound level, articulation, and timbre.

*Computer-Teacher condition*. The computer feedback was transferred into hand-written notes (similar to the Teacher-Teacher condition). In this condition, we used confederates with knowledge about research design who could easily grasp the deception aspects of the study and could act as music teacher. Four colleagues from the department with guitar playing experience were our confederates. They read the feedback suggestions aloud to the performer, handed over the paper with feedback, and left the laboratory.

Neither the performers nor the music teachers knew beforehand the real purpose of the experiment. Some time after the data had been collected and analyzed, the performers and the teachers were carefully debriefed about the true purpose of the study and received a summary of the results.

# Results

*Feedback quality*. The ratings were subjected to a two-way analysis of covariance, featuring two levels of *feedback delivery* (teacher, computer) and two levels of *feedback production* (teacher, computer) as the independent variables and with the background variable *computer attitude* as a covariate. The analysis revealed a significant main effect of *feedback delivery* on rated feedback quality. As predicted, performers judged the quality higher when they believed the feedback derived from a music teacher ($M$ = 8.09, $SD$ = 1.53) than when they believed it derived from a computer ($M$ = 7.34, $SD$ = 1.61,), $F$ (1,75) = 5.16, $p$ < .05, $d$ = .48. In addition, there was a significant main effect of *feedback production*, although contrary to our prediction, performers rated the quality higher when the feedback was produced by a teacher ($M$ = 8.10, $SD$ = 1.22) than when it was produced by a computer ($M$ = 7.33, $SD$ = 1.86), $F$(1,75) = 6.52, $p$ < .05, $d$ = .49. There was no significant interaction between *feedback production* and *feedback delivery*, $F$(1,75) = 0.54, *ns*, but the covariate *computer attitude* yielded a significant effect, $F$(1,75) = 8.53, $p$ < .01. Performers with a positive view towards using computers as learning tools for expression in music tended to rate the feedback quality higher overall than those with a negative view.

*Additional ratings*. The ratings of *feedback comprehensibility* were subjected to a two-way analysis of covariance with the same independent variables as above, but with *learnability attitude* added as the covariate. The results indicated no significant main effects of either *feedback delivery* or *feedback production*, nor was there any significant interaction. In other words, the feedback was judged as equally easy to comprehend regardless of the condition. The covariate yielded a significant effect: Performers with a positive attitude towards the possibility of learning to express emotions in music performance tended to rate *feedback comprehensibility* higher overall than those with a negative attitude.

The ratings of *feedback detail level* were subjected to a two-way analysis of covariance, with the same independent variables as above, although with *musical experience* included as the covariate. The was no significant main effect of *feedback delivery*, but there was a significant effect of *feedback production*. More specifically, performers rated the feedback produced by a teacher ($M$ = 8.55, $SD$ = 1.46) as more detailed than the feedback produced by a computer ($M$ = 6.21, $SD$ = 2.61, $d$ = 1.10), $F$(1,75) = 21.25, $p$ < .01. There was no significant interaction between *feedback delivery* and *feedback production*, and the covariate failed to reach significance.

*Free comments*. There was also room for the performers to write free comments to their feedback ratings. Comments on quality and detail level were categorized by two independent coders, and inter-coder agreement was estimated using Cohens Kappa (Howell, 1992). Mean inter-coder agreement was $\kappa = .85$ for feedback quality and $\kappa = .83$ for feedback detail level. To summarize the results, for the feedback quality comments (categorized into *positive*, *negative*, *both positive and negative*, and *miscellaneous* comments) the largest proportion of positive comments was observed in the Teacher-Teacher condition ('he knew what he was talking about'). The proportion of positive comments decreased as a function of computer involvement, with the largest proportion of negative comments ('too simple') observed in the Computer-Computer condition.

With regard to feedback detail level (comments categorized into *sufficient detail*, *insufficient detail*, and *miscellaneous*), the teacher-produced feedback received more comments implying a sufficient detail level. One tendency in the comments was that performers appreciated teacher-produced feedback because it explained how the feedback suggestions should be realized in the performance, whereas the computer-produced feedback did not (e.g., 'More staccato: How?').

# Conclusions

The results indicate that verbal feedback was judged as being significantly better in quality when it was believed to originate from a teacher, than when it was believed to originate from a computer. This confirms that generally negative attitudes regarding computers influence performers' impressions in ways that may have very little to do with the actual merit of the applications concerned. The results also indicated that performers rated the feedback as higher in quality when it *did* derive from a teacher, regardless of whether it was believed to derive from a teacher or a computer. The teachers' feedback was more elaborate than the computer's feedback. The teachers provided encouragement, examples, and explanations, whereas the computer feedback was concise and focused only on the acoustic cues of primary importance to the communicative process.

Even though Study II showed that feedback from the computer program feedback enhanced performers' emotional expression more efficiently than teacher feedback, the present study suggests that performers prefer teacher-produced feedback. This preference appears to reflect both general attitudes towards computers and preferences regarding the precise feedback contents.

When developing novel methods for teaching expression, these need to be carefully introduced in musical education. Performers need to be informed about the theoretical and empirical bases of these applications. Further, the applications need to be designed so that they are perceived as more human-like during the interaction, for example by using more colorful language, providing pedagogical examples, and offering reasons for the suggestions provided – but without loosing the precision and accuracy that feedback from a computer program could provide (Study II). If these requirements are met, computer-assisted teaching could become an important complement to traditional teaching in music education.

# General discussion

## Main findings

The overall aim of this thesis was to develop and evaluate a new empirically-based approach to teaching emotional expression in music performance. Study I investigated how expression is currently taught in its natural context. One important result was that the focus in the teaching was on technique and notation, but that expression was also considered – mainly in an implicit and vague manner. The findings further indicated that the feedback provided by music teachers to students was primarily outcome feedback (i.e., the teacher provides information about whether the performance was good or bad, but no information about why). It was concluded that there is a need for novel methods for teaching expression and that such methods should offer explicit and informative feedback to the performer.

Study II presented a new and empirically-based method for teaching emotional communication in music performance: the Feel-ME program. The program automatically analyzes a performer's manner of communicating emotions and provides cognitive feedback that makes the performer able to compare his or her performance model with an optimal performance model, including explicit suggestions on how to change certain acoustic features to enhance the expression. Results from two listening tests indicated that those participants who received computer feedback or teacher feedback improved their communication more than those who did not receive feedback, and that those who received computer feedback improved more than those who received teacher feedback. Though the performers in the computer feedback group rated the program as easy to use and understand, they indicated that they did not want to use it in their own learning of expression.

Could the performers' negative impressions of the Feel-ME program in Study II be due to negative attitudes towards computers or a dislike of the characteristics of the actual feedback contents? In Study III, we manipulated (a) the participants' belief about whether the feedback was produced by a teacher or a computer program (feedback delivery) and (b) the feedback content in terms of whether it was actually produced by a teacher or a computer program (feedback production). The results showed that the mere belief that the feedback derived from a teacher yielded higher quality ratings, but so did also feedback that *did* indeed derive from a teacher. Thus, both feedback contents and negative attitudes towards computers could account for the negative impressions of computer-assisted teaching of expression in Study II and in previous studies. This is an important result that should be taken into account when developing and introducing computer applications for teaching of emotional expression in music performance.

## Limitations and methodological issues

There are several limitations in the present thesis that should be mentioned. First, the relatively small number of cases investigated in Study I do not allow generalization. Some results may be limited to individual lessons or to the Swedish context of the study. Even so, the results add to the slowly growing body of evidence from similar case studies in a variety of settings (Rostvall & West, 2001; Young, Burwell, & Pickup, 2003). When filming people, there is always a certain risk that the video camera influences their behaviors. Although both teachers and students claimed that the camera did not affect them, the fact remains that during natural conditions, there is no camera present in the teaching situation.

With respect to Study II, the findings need to be replicated using other performers, instruments, melodies, methods, and contexts. For instance, we used only one melody (When the Saints). However, Juslin (2000) found that the cue utilization is more consistent across different melodies than across different performers. This means that the performer uses mostly the same code regardless of the melody, but that there are large individual differences among performers. To obtain generalizable knowledge, it may thus be more important to include many performers than to include many melodies. Still, an important future goal is to evaluate the program in the field to increase the generalizability in terms of instruments, repertoire, and settings; explore the possible long-term benefits; and look into individual differences among performers. The efficacy and usability of the program is likely to depend on the individual user, as well as on the specific context of its use. In Study II, performers were put in a situation where they had to interact with a computer program in a controlled laboratory setting without any information about the program's theoretical background.

As regards the comparison of teacher feedback and computer feedback in Study II, it should be noted that teachers were prevented from using musical modeling (i.e., imitation of a sound model), which could seem to render the condition unrealistic. However, several studies have suggested that relatively little time is devoted to musical modeling during a lesson (see Study I; Sang, 1987; Speer, 1994). Even so, a condition that would have allowed teachers to use musical modeling could have produced different results.

In order to avoid cognitive overload and ceiling effects, feedback sessions in Study II focused only on the two emotions that the performer was least successful in conveying. Focusing on the two lowest and most extreme values brings the risk of statistical regression towards the mean. But this risk was equal in all three conditions and could not explain why the CFB group improved more than the other groups. A more serious problem with the design is that it did not allow for comparisons of the relative efficacy with which the communication of individual emotions were improved.

Regarding Study III and its particular design, neither the performers nor the teachers were allowed to know the real purpose of the study until all data had been collected. The risks of participants acquainted with one another telling each other about the experiment or of performers finding out the idea of the experiment during the procedure were minimized and controlled in various ways. Yet, some participants (6%) knew that there was 'something that you could not pass on to others' and may have had certain expectations on the experiment. This could have had an affect on their behavior during the experiment.

The background questionnaire in Study III containing questions regarding attitudes towards computers and learning expression was administered at the end of the procedure in order to not reveal anything about the real purpose of the experiment. In that respect, it was not an optimal 'background' measure, particularly for participants who believed they received computer feedback. However, we obtained the measures after the ratings of feedback quality, so that questions about computers would not raise suspicions about computers being involved in the experiment, thus influencing the feedback ratings.

## The Feel-ME program

There are a number of advantages and disadvantages of the present approach to learning expression, regarding computer-assisted teaching in general and the Feel-ME program in particular.

Advantages of a computer-assisted teaching approach, such as the Feel-ME program, include having unlimited time, providing feedback in a relaxed social environment, and providing possibilities for flexible and individually-based learning. The Feel-ME program, in particular, may provide objective and specific feedback by explicitly describing the relationships between expressive intentions, acoustic cues, and listener impressions – relations that normally lie hidden in tacit knowledge. Although one could fear that use of the Feel-ME program could lead to a standardization of performances of music, it must be noted that the decision about how to interpret the music is left to the performer. In addition, the fact that the communicative process involves several partly redundant cues means that there is much flexibility in the communicative process: different cues may substitute for one another

(Juslin, 2001). Besides being a useful practice tool, the Feel-ME program could serve as a diagnostic test of expressive skills and assist music teachers in identifying weaknesses in a performer's expressive strategy. Finally, the Feel-ME program might also be useful as a research tool by helping music researchers to quickly analyze large samples of performances.

The disadvantages of computer-assisted teaching include dependency on computers and electricity, that only certain aspects of the performance are addressed, and 'non-sensitive' feedback. In its current form, the Feel-ME program can only analyze acoustic cues from brief extracts of monophonic performances (i.e., melodies). It is thus mainly suitable for solo instruments, practicing pieces in short segments. Also, it does not take into account local expressive variations that could be important in expression of emotions, or visual aspects of a performance, such as body language, gestures, or facial expressions.

The Feel-ME program should be regarded as a complement in music education. Whereas humans are creative and good at interpreting ambiguous situations, machines are good at precise and reliable operation. In that way, computer-assisted teaching, such as the Feel-ME program, could contribute precise feedback about emotional expressions, whereas music teachers could provide knowledge on how to 'break the rules', help to find a balance among different aspects of the performance, and boost the performer's confidence.

## Implications and future directions

In this thesis, it has been demonstrated that it is possible to learn to play expressively, given that informative feedback is provided. The Feel-ME program was effective in improving the accuracy of the communicative process. Performers were able to make use of explicit feedback regarding individual acoustic cues and to translate such patterns into altered patterns of playing. Hence, contrary to previous claims that musical expression is "too subjective and individualistic for measurement" (Hoffren, 1964, p. 32), it is really possible to measure and predict most aspects of musical expression. If performers employ acoustic cues in certain ways in order to communicate an intention (e.g., 'happiness' using fast tempo, moderate to loud sound level, staccato articulation, and bright timbre), listeners will be able to recognize the expression. The applied research presented in this thesis, based on the GERMS model and the lens model, helps 'demystify' expression and shows that it is not merely a subjective quality that cannot be explained in scientific terms.

The Feel-ME program has been shown to work in the laboratory (i.e., music performers improve their communication of emotions), and it would thus be interesting to see if the program could work in the field. By using the program in the field during a longer period, we would be able to study long-term effects of interacting with it, how learning changes over time, and also how impressions of the program changes over time. The Feel-ME program as a naturally occurring feature at music conservatoires would perhaps help to change performers' attitudes in a positive direction.

Music performers receive and evaluate novel computerized methods both depending on general attitudes towards computers and on the nature of the feedback as such. Successful implementation of computer-assisted teaching should thus somehow address attitudes, for instance by providing sufficient information. Negative attitudes might be off-set by carefully explaining the theoretical bases of the applications and how they ultimately reflect human knowledge rather than arbitrary results from a computer.

Teacher-produced feedback was rated as better than computer- produced feedback in Study III because it was richer in information and provided concrete examples on how to achieve various effects – often complemented by explanations of why the proposed changes would 'work'. It might thus useful to combine specific computer feedback with explanatory examples. However, designing the feedback suggestions so that the feedback remains specific without containing too much information is a delicate matter that needs careful consideration in future research and development.

When feedback was delivered by a computer (and in some cases also by a teacher), some performers in Study III were not too keen on the idea of being presented with a 'mould' or a 'standard' way of expressing emotions (e.g., "Very exact, but I am doubtful whether there really are exact definitions of happy music" and "Isn't it individual how you imagine an emotion – must happiness be fast?"). It is important here to distinguish a connotative level of expression and a more personal level of expression. The Feel-ME program provides feedback that is based on models of how people typically judge or perceive emotions in music performances intended to communicate specific emotions. The focus is on expression as part of a communicative system, a connotative level of expression (in which performers and listeners agree on how certain emotions are expressed and recognized). If, on the other hand, a performer is only playing a piece of music to 'vent' his or her own emotions without any regard to whether these emotions are correctly identified or not by other individuals, this is an example of expression on a personal level. However, a common strategy for many professional musicians is to rely on connotative expression, and to 'add on' a personal expression, so that, for instance, happiness is clearly communicated to an audience, although with a personal touch.

Another interesting issue would be to expand the lens model in future implementations. In a study by Juslin and Lindström (2003), the lens model was expanded with respect to how both composed and performed features contributed to the expression of a piece of music. It would be interesting to expand or complement the lens model in several other ways, such as adding contextual cues (e.g., real-time audience reactions such as cheering, booing, coughing) or visual cues (e.g., having a webcam reading facial expressions or gestures). It would also be useful to have stored data on how a performer usually changes his or her cues when improving certain emotion expressions, which could be included in the feedback if the improvement is minimal even after several feedback cycles in the Feel-ME program.

The present thesis focuses on the emotional expression component of music performance. However, emotional expression should not be the only goal of practice. Performers must develop other aspects as well, using other means (e.g., Juslin, 2003). An important question for further research is how different teaching strategies might be effectively combined in an overall practice approach. For instance, it would be interesting to explore how other applications of computer-assisted teaching that provide feedback can be used in combination with the Feel-ME program. Examples of computer feedback that focuses on other aspects of performance can be found in projects such as IMUTUS (Interactive Music Tuition System, http://www.exodus.gr/imutus/) and VEMUS (Virtual European Music School, http://www.vemus.org/), and in music programs such as PlayPro (where one learns to play the guitar with the help of tablature) and Smartmusic (a plug-in to Finale).

## Concluding remarks

Expression is not addressed in a satisfactory manner in current instrumental teaching, and reasons for this may be found both within the field of music education itself and in earlier research on performance expression. Teaching of expression appears to be influenced by myths, whereas researchers have tended to treat expression as a single category. The present thesis presents a new empirically-based approach to expression based on the GERMS model, in which expression is conceptualized as a multi-dimensional phenomenon. The communicative process of emotional expression is explicitly realized in terms of a modified lens model. The present thesis suggests that a computer program that automatically analyzes acoustic cues of music performances, creates models of playing strategies, and provides informative feedback to performers can improve their communication of emotions. Both the present thesis and other studies that have compared computer-assisted teaching with traditional teaching suggest that computer-assisted teaching can be effective (Webster, 2002). However, computer-assisted teaching of expression needs to be carefully introduced in order for students and teachers to be willing to adopt such an approach.

# Summary in Swedish

**En ny metod för undervisning i känslomässigt uttryck i musikutförande**

En av de viktigaste aspekterna av musikutförande är förmågan att uttrycka känslor. Ändå tyder forskning på att musikundervisning försummar denna aspekt. Syftet med denna avhandling var att presentera och utvärdera en ny metod för undervisning i känslomässigt uttryck i musikutförande.

Studie I undersökte hur musikundervisning går till i sin naturliga miljö, med särskilt fokus på känslomässigt uttryck. Trots att det fanns skillnader mellan lärare utmärktes samtligas undervisning av en brist på tydliga mål, specifika uppgifter och informativ feedback.

Studie II presenterade och utvärderade ett nytt datorprogram som analyserar musikutföranden för att ge informativ feedback med specifika förslag på hur det känslomässiga uttrycket kan förbättras. Musiker delades slumpmässigt in i tre grupper: (1) feedback från datorprogrammet, (2) feedback från en lärare samt (3) repetition utan feedback. Lyssningsexperiment visade att den största förbättringen ägde rum i den grupp som fick feedback från datorprogrammet. Musikerna tyckte att programmet var lätt att förstå och lätt att använda, men ville ändå inte använda det i sitt framtida lärande.

Studie III undersökte om oviljan att använda det nya datorprogrammet kunde bero på generellt negativa attityder till datorer eller på att man ogillade själva innehållet i datorns feedback. Resultat från ett experiment med det egentliga syftet dolt för deltagarna visade att feedback bedömdes som bättre om man *trodde* att den härrörde från en lärare än om man trodde att den härrörde från en dator. Resultaten visade också att feedback bedömdes som bättre när den verkligen härrörde från en lärare. Musikerna uppskattade i synnerhet lärares exempel och förklaringar till varför deras förslag skulle fungera.

Avhandlingen visar att musiker kan förbättra förmågan att uttrycka känslor genom datorstödd undervisning, men visar också att sådan undervisning kan vinna på att inkludera aspekter som uppmuntran, exempel och förklaringar för att göra metoden mer attraktiv för potentiella användare.

# Acknowledgements

Accomplishing this thesis would not have been possible if it hadn't been for a bunch of people helping out and supporting my work. So, here go the thanks…

First and foremost, I am grateful to my supervisor **Patrik** for making this thesis possible by admitting me as his PhD student and always helping me out with whatever thing (be it statistics or design discussions, pep talks, or cocktail tips). Thanks for being open minded, positive, and having a lot of patience.

My colleagues and friends at instn: **Åsa**, my dear friend and room mate – thanks for all the love; **Hällena**, my dear friend (you're great) – thanks for all the belief you have in me; **Hanna** – you rock, I have never seen a better looking "Death" than you; **Vanda**, my dear friend with a Jacuzzi – obrigada and se faz favor; **Simon** – "asmaxat"; **Malin** – you're the best; **Luisa** – I hope you do get a stone house in Portugal someday; **Thomas Å** – I want to take one of your music quizzes; **Olga** – thanks for sharing the love of cats and music; **Helena Ö**, **Pär**, **Claes**, **Kerstin, Dorota** and the rest of the baby-lab army – thanks for all the support; the music/dance people at the department **Petri L**, **Alf**, **Ingrid**, **Erik** (who is at BTH nowadays) and the KTH guys **Erwin**, **Anders**, and **Roberto** – nice working with you all; and **Sari** – thanks for discussions and support.

I will miss the heavy metal evenings, whisky/skumpa/glögg evenings, and the recurrent farewell parties for Vanda at instn!

Special thanks to: my parents **Kent** and **Anita** for always being there for me and for being who they are; my sister **Jennie** for always being there and for always filling out the position as bass player in the old punk rock band when needed; my grandmother **Doris** for always being there and for all the 'sassa' (dansa) in the past; **Loffe** for understanding the need to cuddle; and the ones who did not make it this far, my grandparents **Sixten** and **Thea** from Williamsro, and **Sune** from Svedet – you are greatly missed and will always have a special place in my heart.

Last, but most importantly, **Petri**, you bring meaning to my life. Without you I am truly lost.

Uppsala, January 2008
Jessika Karlsson

# References


Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106*, 410-433.

Baroni, M., Caterina, R., Regazzi, F., & Zanarini, G. (1997). Emotional aspects of singing voice. In A. Gabrielsson (Ed.) *Proceedings of the Third Triennial ESCOM Conference* (pp. 484-489). Uppsala, Sweden: Department of Psychology, Uppsala University.

Barten, S. S. (1998). Speaking of music: The use of motor-affective metaphors in music instruction. *Journal of Aesthetic Education, 32*, 89-97.

Bartle, B. K (1987). *Computer software in music and music education: A guide*. Metuchen, NJ: Scarecrow.

Boyd, J., & George-Warren, H. (1992). *Musicians In Tune: Seventy-Five Contemporary Musicians Discuss the Creative Process*. New York: Fireside.

Brehmer, B. (1994). The psychology of linear judgment models. *Acta Psychologica, 87,* 137-154.

Brunswik, E. (1956). *Perception and the representative design of experiments*. Berkeley, CA: University of California Press.

Clarke, E. F. (1988). Generative principles in music performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1-26). Oxford, UK: Clarendon Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Davies, S. (1994). *Musical meaning and expression*. Ithaca, NY: Cornell University Press.

Denzin, N. K., & Lincoln, Y. S. (2000). Introduction: the discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.) (pp. 1-28). London: Sage.

Dickey, M. R. (1992). A review of research on modeling in music teaching and learning. *Bulletin of the Council for Research in Music Education, 113*, 27-40.

Duke, R. A. (1999/2000). Measures of instructional effectiveness in music research. *Bulletin of the Council for Research in Music Education, 143,* 1-48.

Ebie, B. D. (2004). The effects of verbal, vocally modeled, kinesthetic, and audio-visual treatment conditions on male and female middle school vocal music students' abilities to expressively sing melodies. *Psychology of Music, 32*, 405-417.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Friberg, A, Schoonderwaldt. E., & Juslin, P. N. (2007). CUEX: An algorithm for automatic extraction of expressive tone parameters in music performance from acoustic signals. *Acta Acustica united with Acustica, 93,* 411-420.

Gabrielsson, A. (1987). Once again: The theme from Mozart's piano sonata in A major (K. 331): A comparison of five performacnes. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 81-103). Stockholm: Publications issued by the Royal Swedish Academy of Music, No. 55.
Gabrielsson, A., (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of Music* (2nd ed.) (pp. 501-602). San Diego, CA: Academic Press.
Gabrielsson, A., (2002). Emotion perceived and emotion felt: same or different? *Musicae Scientiae, special issue 2001-2002*, 123-147.
Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music, 3,* 221-272.
Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music, 24*, 68-91.
Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.) *Handbook of affective sciences* (pp.503-534). New York: Oxford University Press.
Gabrielsson, A., & Lindström, E. (1995). Emotional expression in synthesizer and sentograph performance. *Psychomusicology, 14,* 94-116.
Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological Review, 108,* 33-56.
Goolsby, T. W. (1997). Verbal instruction in instrumental rehearsals: A comparison of three career levels and preservice teachers. *Journal of Research in Music Education, 45,* 21-40.
Hallam, S. (1997). What do we know about practicing? Towards a model synthesizing the research literature. In H. Jørgensen & A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice.* Oslo: Norges musikhögskola.
Hallam, S. (1998). *Instrumental teaching: A practical guide to better teaching and learning.* Oxford: Heinemann.
Hammond, K. R. (1971). Computer graphics as an aid to learning. *Science, 172*, 903-908.
Hepler, L. E. (1986). *The measurement of teacher/student interaction in private music lessons and its relation to teacher field dependence/field independence.* Unpublished doctoral dissertation, Case Western Reserve University.
Hoffren, J. (1964). A test of musical expression. *Council for Research in Music Education, 2*, 32-35.
Howell, D. C. (1992). *Statistical methods for psychology* (3rd ed.). Belmont, CA: Duxbury Press.
Juslin, P. N. (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception, 14*, 383-418.
Juslin, P. N. (1998). A functionalist perspective on emotional communication in music performance. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 78.* Uppsala, Sweden: Uppsala University Library.
Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1797-1813.
Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin, & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 309-337). New York: Oxford University Press.

Juslin, P. N. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music, 31*, 273-302.

Juslin, P. N., Friberg, A., & Bresin, R. (2002). Toward a computational model of expression in music performance: The GERM model. *Musicæ Scientiæ, Special Issue 2001-2002*, 63-122.

Juslin, P. N., Friberg, A., Schoonderwaldt, E., & Karlsson, J. (2004). Feedback-learning of musical expressivity. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques for enhancing performance* (pp. 247-270). New York: Oxford University press.

Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicæ Scientiæ, 4*, 151-183.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion, 1*, 381-412.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770-814.

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research, 33*, 217-238.

Juslin, P. N., & Lindström, E. (2003). *Musical expression of emotions: Modeling composed and performed features.* Paper presented at 5th Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Hanover, Germany, September 2003.

Juslin, P. N. & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception, 17*, 197-221.

Juslin, P. N., & Sloboda, J. A. (Eds.). (2001). *Music and emotion: Theory and research*. New York: Oxford University Press.

Juslin, P. N., & Västfjäll, D. (in press). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences.*

Jørgensen, H. (2000). Student learning in higher instrumental education: who is responsible? *British Journal of Music Education, 17,* 67-77.

Konishi, T., Imaizumi, S., & Niimi, S. (2000). Vibrato and emotion in singing voice (abstract). In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.) *Proceedings of the Sixth International Conference on Music Perception and Cognition, August 2000* (CD rom). Keele, UK: Keele University.

Kostka, M. J. (1984). An investigation of reinforcements, time use, and student attentiveness in piano lessons. *Journal of Research in Music Education, 32,* 113-122.

Kotlyar, G. M. & Morozov, V. P. (1976). Acoustic correlates of the emotional content of vocalized speech. *Soviet Physics. Acoustics, 22,* 370-376.

Laukka, P. (2004). Instrumental teachers' views on expressivity: A report from music conservatoires. *Music Education Research, 6*, 45-56.

Lindström, E. (1999). *Expression in music: Interaction between performance and melodic structure*. Paper presented at the Meeting of the Society for Music Perception and Cognition, Evanston, USA, 14-17 August 1999.

Lindström, E. (2003). The contribution of immanent and performed accents on emotional expression in short tone sequences. *Journal of New Music Research, 32,* 269-280.

Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research studies in Music Education, 20*, 23-47.
London, J. (2002). Some theories of emotion in music and their implications for research in music psychology. *Musicae Scientiae, Special Issue 2001-2002,* 23-36.
Marchand, D. J. (1975). A study of two approaches to developing expressive performance. *Journal of Research in Music Education, 23*, 14-22.
Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.
McPherson, G. E. (1993). *Factors and abilities influencing the development of visual, aural, and creative performance skills in music and their educational implications.* Unpublished doctoral dissertation, University of Sydney, Sydney, Australia.
Minassian, C., Gayford, C., & Sloboda, J. A. (2003). *Optimal experience in musical performance: a survey of young musicians.* Paper presented at the Meeting of the Society for Education, Music, and Psychology Research, London, March 2003.
Ohgoshi, K. & Hattori, M. (1996b). *Acoustic correlates of the emotional expression in vocal performance*. Paper presented at the Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii, 2-6 December 1996.
Palmer, C. (1996). Anatomy of a performance: Sources of musical expression. *Musical Perception, 13*, 433-453.
Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115-138.
Persson, R. S. (1993). *The subjectivity of musical performance: An exploratory music-psychological real world enquiry into the determinants and education of musical reality.* Doctoral dissertation, University of Huddersfield, Huddersfield, UK.
Persson, R. S. (1996). Concert musicians as teachers: On good intentions falling short. In A. J. Cropley, & D. Dehn (Eds.), *Fostering the growth of high ability: European perspectives* (pp. 303-320). Norwood, NJ: Ablex.
Rapoport, E. (1996). Emotional expression code in opera and lied singing. *Journal of New Music Research, 25*, 109-149.
Rostvall, A-L., & West, T. (2001). Interaktion och kunskapsutveckling [Interaction and learning. A study of music instrument teaching]. Unpublished doctoral dissertation, KMH Förlaget, Stockholm.
Ryan, G. W, & Bernard, H. R. (2000). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.) (pp. 769-802). London: Sage.
Sang, R. C. (1987). A study of the relationship between instrumental music teachers' modeling skills and pupil performance behaviors. *Bulletin of the Council for Research in Music Education, 91,* 155-159.
Seashore, C. E. (1938). *The psychology of music.* New York: McGraw-Hill.
Shove, P., & Repp, B. H. (1995). Musical motion and performance: theoretical and empirical perspectives. In J. Rink (Ed.), *The practice of performance. Studies in musical interpretation* (pp. 55-83). Cambridge: Cambridge University Press.
Silverman, D. (2000). Analyzing talk and text. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.) (pp. 821-834). London: Sage.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, *35A,* 377-396.

Sloboda, J. A. (1996). The acquisition of musical performance expertise: Deconstructing the 'talent' account of individual differences in musical expressivity. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 107-126). Mahwah, NJ: Erlbaum.

Speer, D. R. (1994). An analysis of sequential patterns of instruction in piano lessons. *Journal of Research in Music Education, 42,* 14-26.

Tait, M. (1992). Teaching strategies and styles. In R.Cowell (Ed.), *Handbook of research on music teaching and learning* (pp. 525-534). New York: Schirmer.

Vosskuhler, B. (2005). Musical expression and ways to teach it. *Music Staff: Teacher Lounge, 24*, 1-4.

Webster, P. R. (2002). Computer-based technology and music teaching and learning. In R. Colwell, & C. Richardson (Eds.), *The new handbook of research on music teaching and learning* (pp. 416-439). New York: Oxford University Press

Widmer, G., & Goebl, W. (2004). Computational models of expressive music performance: the state of the art. *Journal of New Music Research, 33*, 203-216.

Woody, R. H. (1999). The relationship between explicit planning and expressive performance of dynamic variations in an aural modeling task. *Journal of Research in Music Education, 47*, 331-342.

Woody, R. H. (2000). Learning expressivity in music: An exploratory study. *Research Studies in Music Education, 14*, 14-23.

Yarbrough, C., & Price, H. E. (1989). Sequential patterns of instruction in music. *Journal of Research in Music Education, 37,* 179-187.

Young, V., Burwell, K., & Pickup, D. (2003). Areas of study and teaching strategies in instrumental teaching: a case study research project. *Music Education Research, 5*, 139-155.

# Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 36

Editor: The Dean of the Faculty of Social Sciences


A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences".)


Distribution: publications.uu.se
urn:nbn:se:uu:diva-8456

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2008