9th International Conference on Music Perception and Cognition
Alma Mater Studiorum University of Bologna, August 22-26 2006

Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience

Mary Broughton
MARCS Auditory Laboratories,
University of Western Sydney,
NSW, Australia
email
m.broughton@uws.edu.au

Kate Stevens
MARCS Auditory Laboratories,
University of Western Sydney,
NSW, Australia

Stephen Malloch
MARCS Auditory Laboratories,
University of Western Sydney,
NSW, Australia

ABSTRACT

Research in many fields has demonstrated the perceptual advantages of experiencing the world through multiple sensory modalities for accurate and effective communication. The aim of the current study was to test the assumption that visual perception of movement plays a role in communicating a musically expressive performance. In the live, concert music setting, performers have increased opportunities for engaging audience attention and guiding awareness to musical content, by presenting information simultaneously via multiple modalities. Non-verbal behaviours and gestures are natural and integral components in interpersonal communication. This study is concerned with investigating the interaction of auditory and visual information in communicating musical expression to an audience. This study is of particular relevance to the marimba (a tuned, wood instrument in the percussion family) because of its relatively restricted range of expressive capabilities such as articulation and duration.

It was hypothesised that multi-modal (audio-visual) perception, where the visual features are expressive and reinforce the performer's expressive musical intention (aural features), would enhance the observer's level of interest and perceived expressivity relative to auditory only perception.

Musically expert and novice observers rated digitised presentations of solo marimba excerpts (projected or deadpan performance manners) on rating scales under two conditions: audio alone and combined audio-visual. The experimental design consisted of three factors each with two levels: modality (auditory alone; combined auditory and visual conditions), stimulus (projected performance manner; deadpan performance manner) and expertise of observer (musically trained; non-musically trained) with the first two variables as repeated measures. The dependent variables were observers' ratings of interest and perceived expressivity indicated on two separate seven-point Likert scales. The marimba was used as the instrument to create digital stimulus materials as the movements required to play it are visible and its inherent expressive capabilities are relatively limited. The stimulus material comprised sets of thirty-two 20-25 second excerpts of 20th century solo marimba repertoire of fast and slow tempi and varying levels of difficulty and musical style, performed by two professional marimbists, one male and one female.

Results support the assumed perceptual advantage of experiencing a musical performance through complementary multiple sensory modalities. Observers could discriminate between expressive and inexpressive performances in both an audio only and audio-visual condition. Observers could most effectively differentiate

In: M. Baroni, A. R. Addessi, R. Caterina, M. Costa (2006) Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9), Bologna/Italy, August 22-26 2006.

ISBN 88-7395-155-4

between an expressive and an inexpressive performance when the presentation was audio-visual. Musically trained participants recorded higher ratings than their non-musically trained counterparts.

INTRODUCTION

Music is a form of high-level motor performance, but with an added dimension of heightened nonverbal communication. In the current climate where music is most readily available in an audio only format (CD, MP3), we ask the question, "Is there a perceptual advantage to seeing and hearing a musical performance?".

Multimodal Perception

When observing a live music performance, an audience member is inundated with a wealth of aural, visual, spatial and movement stimuli. While the focal stimulus is aural (including pitch, timbre, rhythm, form, dynamics and articulation), in live music performance, both aural and visual modes are integrated, influencing the audience member's aesthetic experience (McClaren 1988).

Research has shown that there are many perceptual advantages in experiencing the world through multiple sensory modalities. Massaro (1987) has conducted much experimental research in audio-visual speech integration demonstrating that the combination of auditory and visual stimuli contribute to the most effective communication of content. When the same message is sent simultaneously through more than one mode, redundancies occur, resulting in more accurate communication (Wickens, Lee et al. 2004).

In a recent study investigating sensory integration and perception of musical performance, Vines et al. (in press) found that auditory and visual information interacted, enhancing communication of content. This highlights the perceptual advantages of both seeing and hearing a musician perform.

Non-verbal Behaviour

Non-verbal behaviour as a function of communication and interpersonal interaction, and its relation to spoken language, has been the focus of much research in the past. According to Argyle (1988), nonverbal communication can express what cannot be put into words. He identifies five types of bodily communication: expressing emotions (face, body, voice); communicating interpersonal attitudes; accompanying and supporting speech (auditory mode); self-presentation (for example, appearance;) and rituals (for example, greetings). Of these, expressing emotions and accompanying and supporting speech, are most relevant to musical communication.

The language of music shares many syntactical similarities with spoken language. For example, the phrase, whether it be musical or linguistic, forms a salient unit (Aiello 1994). According to Knapp and Hall (2002), speech phrases and movement phrases appear to be closely, and rhythmically coordinated.

Body movement and gesturing is believed to occur naturally in spoken language, and aid in both the language forming processes and in communicating intention. Speech-related gestures that occur in synchrony with and support a stream of spoken language have been found to increase comprehension by "…vivifying ideas…intensifying point…maintaining listener attention and focus and…marking the organizational structure of the discourse." Where complementary verbal and non-verbal behaviours occur, the message being communicated is perceived more accurately by the listener (Knapp & Hall 2002). Dobrogaev (1931, cited Knapp & Hall 2002) found that when natural gestures were inhibited, speech became more laboured and stresses, intonation and expressiveness were reduced.

Just as people gesture when speaking, musicians naturally tend to move expressively as they perform. In this study, we investigate whether observers are sensitive to changes in audio and visual expression across auditory only and auditory-visual conditions.

While other research has sought to quantify, analyze and describe musicians' expressive body movements (Wanderley, Vines et al. 2005; Davidson 2001), that is not the focus of this paper.

Musical Expression

Juslin (2003) defines expression as "a set of perceptual qualities that reflect psychophysical relationships between 'objective' properties of the music and 'subjective' properties of the listener." (p. 276). Musical expression enables the communication of musical meanings to an audience. The score is the means by which composers record their musical thinking. The composer is reliant on the performer to convey the full meaning of the score (Hill 2002).

Whilst the musical score provides a wealth of information to the performer with regard to the composer's expressive intentions, it is not completely prescriptive. The score usually provides the performer with some information intended to aid in interpretation. Sloboda (1996) points out that expressive deviations from an exact rendition of the score are intentional, take account of the musical structure of the score and are detectable by listeners. Relating the performer's subjective experiences to objective features of a piece is vital in playing expressively (Juslin & Persson 2002). For instance, Gabrielsson and Juslin (1996) analysed the way in which the performer manipulated the musical features (tempo, dynamics, timing) to express varied meanings that were confirmed by listeners.

ISBN 88-7395-155-4

There are many musical elements, such as dynamics and articulations, usually notated in the musical score that are vital in the production of a musically expressive performance. In addition to the notated expressive markings that the performer must interpret in sound, there are other factors, such as subtle variations in timing that contribute to an expressive performance. Much research has investigated timing as a function of a musically expressive performance (Todd 1992; Schaffer 1995; Repp 1999; Friberg & Battel 2002). Todd (1992) proposed a model of musical expression that encompassed both timing and dynamics. Deliberate variation in timing, both within the phrase structure and across the broader musical structure, is a salient facet of a musically expressive performance. For instance, timing patterns that followed the temporal course of objects obeying gravitational forces seem to result in a "natural-sounding" performance (Todd, 1995 cited Clarke 2002).

It is the way each individual performer uses timing, dynamics and articulations in the formulation of their musical/artistic image that constitutes an expressive performance (Clarke 2002). Dahl (2000) concurs that it is the variety in timbre, dynamics, duration and articulations that lead to an expressive performance.

The marimba does not have the same expressive capabilities, such as articulation, timbre and duration that other instruments do. Dahl (2000; Dahl & Friberg 2004) notes that the timbres of percussion instruments are not easily changed. Changes in dynamic level and duration between notes are the primary means of expression (Dahl 2000). This would perhaps make the visual aspect of marimba performance (performer's expressive movement) highly important in communicating an expressive performance to an audience.

Investigations have been conducted to discover whether observers were sensitive to performers' emotional expressive intentions in audio only, visual only and combined audio-visual modes. While emotional content is not the focus of this study, it is interesting to note that listeners (audio only) could generally correctly decode performers' intentions with regard to emotional expression/character in violin, flute, electric guitar and singing performances (Gabrielsson & Julsin 1996). Dahl and Friberg (2004) found that observers were, for the most part, able to correctly identify a marimba player's different emotional expressive intentions when viewing (audio-visual) video recordings of the performances showing different parts of the performer's body. In a study investigating spectators' impression (visual only) of emotional expression/character and dancers' intended emotional expression, Camurri et al. (2003) showed that observers were able to detect dancers' different intended emotional expression from performances of the same choreography. These studies indicate that audience members perceive the presence of, and variations in artistic expression via multiple modalities.

Novelty has been found to contribute to instant enjoyment and situational interest. According to Hidi and Anderson's (1992 cited Chen, Darst et al. 2001) definition, situational interest is, "…the appealing effect of an activity…on an individual, rather than the individual's personal preference for the activity." (Chen, Darst et al. 2001 p. 384). It is predicted that an expressive audio-visual performance will be more interesting for an observer, when compared with an inexpressive performance, due to variety in the visual information presented.

Visual Perception of Movement in Music Performance

Visual information can aid an audience in perceiving and understanding the performer's expressive intention (Davidson 1993). Therefore, the body plays an important role, not only the physicalities of playing the instrument, but also in communicating expressive intention to an audience (Clarke 2002).

Davidson (1993) showed the important role visual information played in conveying expressivity in music performance manner. In this study, Davidson conducted two experiments investigating the perceptual information contained in the body movements of violin and piano performers. Four final-year undergraduate violinists performed excerpts from four different pieces from the Baroque, Classical and Romantic violin repertoire. The student pianist performed a selection from a piece by Mussorgsky. Video recordings were made of performances by each musician in three differing performance manners: deadpan (with minimal expressive interpretation of the music), projected (consistent with public performance) and exaggerated (overstating all aspects of the expressive features) that were actually performable. She stated that the deadpan and exaggerated manners were typically used in teaching, while the projected manner was used in public performance. These recordings were made using point-light technique (Johansson 1973). Fifty-five music students, who served as observers, were presented with excerpts, between thirty and seventy seconds in duration, in three modes: sound only, sound and vision and vision only. Findings indicated that there was agreement between performer intention and audience detection of performance manner in all three conditions. Davidson concluded that vision alone seemed to provide more information as to expressive intention. It is unknown whether there was an effect of training on the results as only music students performed ratings. As the stimulus material consisted of different pieces of repertoire performed by different performers, it is not known whether this could have exerted at effect on results.

In an unpublished empirical follow-up study, Davidson (1995) used the stimuli and experimental design from her

1993 study to examine perceptual responses between musicians and non-musicians. Though sample size was small (ten musicians and non-musicians), previous results for the musicians group were confirmed. Non-musicians' performance indicated that vision may be the most reliable means for discrimination between performance manners for this group.

McClaren (1988) studied solo marimba performance examining the effects of performers' visual attributes (body movements) on listeners' perceived quality assessment. Stimulus material was selected from performances of a 20th century solo marimba piece by seventeen university students. A panel of six experienced listeners rated the visual and aural attributes of performances on two seven-point bipolar scales. Three negatively and three positively rated performances were selected and presented in audio-visual and audio only conditions to thirty-seven non-music college students. Participants rated all six performances (in random order) in both an audio only and audio-visual condition on seven-point bipolar rating scales as sensitive-insensitive, effective-ineffective, good-bad and positive-negative. Results revealed that, "…listeners will consistently rate viewed performances higher than heard performances, but only if the visual presentation is positive." (p.57). "Negative" performances did not receive significantly different ratings between audio only and audio-visual conditions. McClaren's study also supported the generally accepted idea that the basis of a good musical performance is the high quality of the aural performance, with positive visual attributes enhancing the audience perception of it as a better performance. It should be noted that ratings were performed by non-music college students so an effect of musical training is not known.

A recent exploratory study into the musical significance of advanced clainetists' ancillary (expressive) gestures was conducted by Wanderley et al. (2005). Performances by two clarinetists of an unaccompanied piece by Stravinksy from the solo clarinet concert repertoire, were recorded in various performance manners (immobilized, standard and expressive), similar to those termed by Davidson (1993) as: deadpan, expressive and exaggerated. Recordings were made using a digital video camera and movement tracking technology for a quantified analysis of gesture. Recorded performances were not used in an experiment to confirm the clarinetists' intended performance manner by impartial observers.

In the aforementioned study investigating sensory integration and perception of musical performance by Vines et al. (in press), thirty musically trained observers were randomly assigned to one of three conditions (audio only, visual only, combined audio-visual). In two separate trails, participants' judgements of perceived tension and phrasing were continuously recorded throughout the presentation of recorded clarinet performances executed by two performers in public performance manner. Results showed that sound dominated observers' perceptual experience of tension. However, participants' judgements of tension at significant points in the performances were either enhanced or diminished by the visual component in the audio-visual mode. In addition, the performers' gesturing (visual information) indicated structural information of a piece, such as phrasing, to an observer. Ratings were performed by musically-trained participants only, so an effect of musical training is unknown. According to Vines et al., pilot testing indicated that musicians and non-musicians performed similarly on this task so results could be generalised to both populations. Participant numbers in each condition were quite small (ten per condition).

Audience Expertise

It is assumed that professional performers are knowledgeable experts in their field and that they use their expertise to present an expressive performance to the audience. Typically, concert audience members are active listeners, engaged in the performance. Audience members bring their musical knowledge and experience to the task of listening. This experience may be extensive or minimal. Listeners who have more musical experience and knowledge (experts) perceive a performance differently to novices (Gromko 1993). Expert listeners can perceive more subtle details and more of the performer's intention from the musical sound alone than novices, while novice listeners are more reliant on visual information for their judgements (Davidson 1997). It is anticipated that musically trained observers will be able to assess whether a performance is expressive or inexpressive and expertise will interact with modality.

Aim, Design and Hypotheses

The aim of the current study was to test the assumption that visual perception of movement plays a role in communicating a musically expressive performance to an audience. The factorial experimental design was comprised of three independent variables, each with two levels. The first between-subjects independent variable was level of expertise of observer (musically trained or musically untrained). The remaining within-subjects independent variables were modality (audio only or audio-visual) and performance manner (projected or deadpan). A visual only condition was not included as this study was concerned with exploring whether there was a positive perceptual effect for observers when they could see a musician whilst they heard them perform as opposed to hearing only. Ratings of interest and expressiveness were recorded by participants for each item on two separate seven-point Likert scales (very uninterested-very interested; very inexpressive-very expressive).

It was hypothesized that participants would record higher ratings for pieces presented in the audio-visual condition in comparison to those presented in the audio only condition. Another hypothesis proposed that participants would assign

higher ratings to pieces performed in a projected manner than those performed in a deadpan performance manner. It was predicted that an interaction would occur between modality and performance manner. It was anticipated that participants would assign higher ratings to pieces performed in a projected performance manner, and lower ratings to pieces performed in a deadpan manner, when presented in the audio-visual condition in comparison to the audio only condition. The researchers were interested to discover whether musically trained participants would assign higher ratings to pieces relative to participants without musical training.

METHOD

Participants

A total of 48 participants took part in the study (17 males, mean age 24.94, *SD* 7.09; 31 females, mean age 23.06, *SD* 9.38). The sample was divided into two equal groups of 24 (musically trained and non-musically trained), based on information about each participant's musical experience gathered via questionnaire. Musically trained participants were those who had completed at least six years of formal training in music and were currently active as a performing, teaching or composing musician (17.29 mean years training, *SD* 11.2). Non-musically trained participants had undertaken less than two years of formal music training (0.7 mean years training, *SD* 0.83). Participants were recruited through a convenience strategy from universities in Sydney, National Music Camp for students in Canberra, and music teachers from schools in Canberra. University of Western Sydney Psychology students received course credit for their participation. It was stipulated that participants must have normal or corrected-to-normal vision and normal hearing for inclusion in the study.

Stimuli

Two professional marimba players (one male and one female), dressed in black, performed two excerpts from four pieces of marimba repertoire by four different composers. Excerpts of compositions performed were Movements II and III from *Marimba Dances* by Australian Ross Edwards; Movements I and III from *Suite No.2 for Marimba* by Japanese composer and marimbist Takiyoshi Yoshioka; *Nancy* by Emanuel Sejourne from France and *Merlin* by Andrew Thomas of the United States of America. Each composition differed in style of music and level of difficulty. All compositions were written in standard musical notation. The excerpts selected for performance from each composition differed in tempi (ie. one was slow and the other was fast). The performers played these excerpts in two different performance manners – projected, as in public performance, and deadpan (without projection as in public performance). Performances of the excerpts were recorded at the Old Darlington School, University of Sydney, Australia. A Malletech Stiletto marimba was the instrument played by performers using Encore Nancy Zeltsman series mallets and Mike Balter mallets. Excerpts were recorded on a Panasonic digital video camera (NV-MX300EN/A) with an external Rode NT4 stereo condenser microphone providing sound through a Behringer mixing desk.

From these audio-visual recordings, ninety-six, twenty to twenty-five second selections (clips), were taken including complete phrases. Editing was performed using Adobe Premier Pro 1.5. The audio-visual computer files were converted into wave files using River Past Audio Convertor 6.5. Group normalisation was performed on the wave files using Adobe Audition 2.0 in order to equalise the volume between performers playing the same excerpt. Each normalised wave file was then relinked to its video. In order to control possible confounds such as facial expression, an opaque, rectangular box was created using the background (off-white painted brick wall) and laid across the head movement area in each clip. This opaque rectangular box disguised the face of the performer but did not interfere with the observer's ability to view the whole body of the performer.

Six sets of sixteen (twenty to twenty-five second) clips were created in the audio-visual format, and then in the audio only format, making a total of twelve sets. The clips contained within each set were balanced in terms of gender of performer and performance manner (expressive or deadpan). Each set contained selections from both of the excerpts of all pieces performed. No excerpt set contained performances of the same twenty to twenty-five second clip, by the same performer, in the same performance manner. Individual clips were only included within each set once.

Within the six audio-visual excerpt sets built, each of the sixteen clips was presented twice. Title frames of two seconds in duration were inserted into the timeline to mark the first and second presentation of each clip. Following the second presentation of a clip, a title frame was inserted that contained the instructions to the participant that they had fifteen seconds to record their response before the next clip would begin. Participants were requested to record their ratings of 'interest' and 'expressiveness' by circling the number that best fit the respondent's judgement on two separate seven-point Likert scales (very uninterested–very interested; very inexpressive–very expressive). A gap of one and a half seconds was left between clips and titles on the timeline.

Once completed, each of the six sets was individually imported into the Master Timeline window and the auto-colour correction effect was applied from the effects window. This procedure was conducted in order to eliminate noticeable changes in lighting that had occurred during the recording phase of the stimulus preparation. From the Master Timeline window, each of the six sets was exported as an .avi file. Each audio-visual excerpt set was twenty minutes in duration.

ISBN 88-7395-155-4

In order to create the audio only versions of the audio-visual excerpt sets, each of the six sets was copied into a new timeline window and the audio and visual information in each of the 16 clips were unlinked and the visual information deleted, leaving the audio information intact. This resulted in the titles remaining, as in the audio-visual versions of the excerpt sets, but the participant would see a black screen when the auditory stimulus was presented. One at a time, the six audio only versions of the sets were imported into the Master Timeline window and exported as an .avi file to a Maxtor One Touch II 200GB portable external hard-drive. Each audio only excerpt set was twenty minutes in duration.

Participants were presented an audio only set of sixteen excerpts and a different audio-visual set of excerpts. Both the audio only and the audio-visual excerpt sets were played through Windows Media Player.

Equipment

The stimuli were presented to participants via Windows Media Player on an LG LS70 Express laptop computer with Koss (UR20) headphones.

Procedure

Participants were presented with an information sheet outlining the study and written consent was gained prior to testing. The testing procedure was conducted on an individual basis in a quiet room. Participants were presented with one of the six sets of audio only clips and a different set of clips selected from the six audio-visual sets. No participant received the same set of excerpts in the audio and audio-visual conditions. The order of presentation of audio and audio-visual sets was counterbalanced in both the musically trained and non-musically trained groups. All order permutations of audio only and audio-visual sets received ratings from two different participants in both the musically trained and non-musically trained groups of participants. Every excerpt set in both the audio only and audio-visual conditions was presented in the first and second position twice.

Each audio only or audio-visual clip contained within a set of excerpts was presented twice. After the second viewing, participants were requested to record their responses indicating how interested they were in it, and how expressive they deemed the excerpt to be, by circling a number on each rating scale that best fit their judgement. In the audio only condition, participants were recording their judgements based on the auditory information they received. In the audio-visual condition, participants were recording their responses based on the auditory and the visual information they received. Responses to interest and expressiveness were recorded on two separate seven-point Likert scales (very uninterested–very interested; very inexpressive–very expressive). Participants were instructed that their ratings of interest and expressiveness may or may not be related. Similar or dissimilar responses to excerpts on the two scales were equally valid. Participants were given a minute's break between the presentation of audio and audio-visual sets. Each set was twenty minutes in duration. Upon completion of the testing procedure, background demographic information and information relating to participants' musical training, experience and personal musical taste was gathered via questionnaire. The questionnaire also contained questions relating to personal taste for the sound of the marimba.

RESULTS

Data consisted of expressiveness and interest ratings. These were analysed separately using two, three-way ANOVAs and will be reported separately.

Expressiveness Ratings

It was hypothesized that participants with musical training would assign higher expressiveness ratings to pieces relative to participants without musical training. This effect was observed $F(1,1528)=50.53$, $p=.00$, with mean expressiveness ratings recorded by trained participants of 5.34 compared with 4.82 for untrained. The second hypothesis stated that participants would record higher ratings for pieces presented in the audio-visual condition in comparison to those presented in the audio-only condition. No main effect was observed. A main effect was observed, $F(1,1528)=83.17$, $p=.00$, in support of the hypothesis that participants would assign higher ratings to pieces performed in a projected manner than those performed in a deadpan performance manner. Mean expressiveness ratings recorded for performance manner were 5.4 for projected and 4.76 for deadpan.

It was hypothesized that an interaction would occur between modality and performance manner. It was anticipated that participants would assign higher expressiveness ratings to pieces performed in a projected performance manner, and lower ratings to pieces performed in a deadpan manner presented in the audio-visual condition in comparison to the audio only condition. A significant interaction was observed between modality and performance manner $F(1,1528)=42.78$, $p=.00$ (see Figure 1.). Mean expressiveness ratings in the audio only modality for projected performance manner were 5.61 and 4.78 for deadpan performance manner. In the audio-visual modality, mean expressiveness ratings for projected performance manner were 5.65 and 4.54 for deadpan performance manner.

ISBN 88-7395-155-4

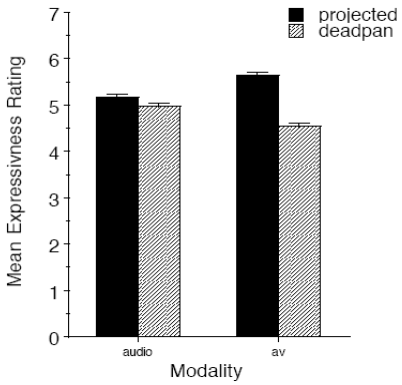

Figure 1. Modality by Performance Manner Interaction Mean Expressiveness Ratings

Interest Ratings

It was hypothesized that participants with musical training would assign higher interest ratings to pieces relative to participants without musical training. This effect was observed $F(1,1528)=50.3$, $p=.00$, with mean interest ratings recorded by trained participants of 5.21 compared with 4.64 for untrained. The second hypothesis stated that participants would record higher ratings for pieces presented in the audio-visual mode in comparison to those presented in the audio-only mode. A main effect was observed $F(1,1528)=4.87$, $p=.027$, with mean interest ratings of 4.84 for the audio-only mode and 5.01 for the audio-visual mode. A main effect was observed, $F(1,1528)=22.28$, $p=.00$, in support of the hypothesis that participants would assign higher ratings to pieces performed in a projected manner than those performed in a deadpan performance manner. Mean interest ratings recorded for performance manner were 5.1 for projected and 4.74 for deadpan.

It was hypothesized that an interaction would occur between modality and performance manner. It was anticipated that participants would assign higher interest ratings to pieces performed in a projected performance manner, and lower ratings to pieces performed in a deadpan manner presented in the audio-visual condition in comparison to the audio only condition. A significant interaction was observed between modality and performance manner $F(1,1528)=11.07$, $p=.001$ (see Figure 2.). Mean interest ratings in the audio-only mode for projected performance manner were 4.89 and 4.78 for deadpan performance manner. In the audio-visual mode, mean interest ratings for projected performance manner were 5.32 and 4.7 for deadpan performance manner.

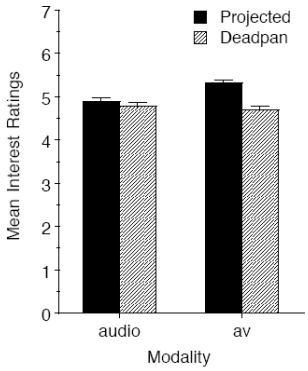

Figure 2. Modality by Performance Manner Interaction Mean Interest Ratings

DISCUSSION

Musically trained participants assigned higher ratings to pieces relative to participants without musical training, supporting the first hypothesis. The second hypothesis, stating that participants would record higher ratings for pieces presented in the audio-visual condition in comparison to those presented in the audio-only condition, was only supported for the dependent variable, interest. Higher ratings were recorded by participants for pieces performed in a projected manner compared with those pieces performed in a deadpan performance manner, supporting the third hypothesis. The predicted interactions between modality and performance manner were supported by results for both dependent variables.

Results indicate that it is more interesting to both see and hear a musician perform. These findings support the assumption that there are perceptual advantages to experiencing a musical performance through complementary multiple sensory modalities (Vines, Krumhansl et al. in press). It has been demonstrated empirically that observers are sensitive to changes in audio and visual expression across auditory only and audio-visual conditions. This supports the notion that as in speech, inhibiting expressive non-verbal behaviours (body movement and gesturing) impacts negatively on sound production, expression and communication of meaning. Also highlighted, is body movement functioning as both instrumental technique and a means to communicate expressivity through both vision and sound (Clarke 2002).

Although a main effect was only observed for interest ratings in the audio-visual mode, the interaction that occurred between modality and performance manner demonstrated observers can most effectively differentiate between an expressive and an inexpressive performance when the presentation is audio-visual. This provides support for the concept that variety in the presentation of visual information is more interesting to an observer (Chen, Darst et al. 2001). In addition, expressive movement provides cues to an audience as to a musically expressive performance (e.g. Davidson 1993). As the repertoire

ISBN 88-7395-155-4

performed for construction of the stimulus material was art music from the 20th century, knowledge and experience with the music of this period could account for higher ratings recorded by the musically trained group (Gromko 1993). It would be interesting to replicate the experiment with repertoire of other periods of music. It would also be interesting to replicate this experiment using other instruments that have more expressive capabilities than individual percussion instruments to see whether results can be generalized to all instrumentalists. Future research could address: a system of analysis of the items that scored most favourably; and development of methods for training advanced music students in expressive and communicative performance skills.

REFERENCES

Aiello, R. (1994). Music and language: Parallels and contrasts. In R. Aiello (Ed.), *Musical perceptions*. New York: Oxford University Press.

Argyle, M. (1988). *Bodily Communication*. Madison, Connecticut, International Universities Press, Inc.

Camurri, A., Lagerlof, I. & Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International journal of human-computer studies, 59*, 213-225.

Chen, A., Darst, P. W. & Pangrazi, R.P. (2001). An examination of situational interest and its sources. *British journal of educational psychology. 71,* 383-400.

Clarke, E. (2002). Understanding the psychology of performance. In J. Rink (Ed.), *Musical performance: A guide to understanding* (pp. 59-72). Cambridge, UK: Cambridge University Press.

Dahl, S. (2000). The playing of an accent: Preliminary observations from temporal and kinematic analysis of percussionists. *Journal of new music research 29,* 225-233.

Dahl, S. & Friberg, A. (2004). Expressiveness of a marimba player's body movements. *Speech, music and hearing quarterly progress and status report 46*, 75-86.

Davidson, J. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of music. 21,* 103-113.

Davidson, J. (1995). What does the visual information contained in music performances offer the observer? Some preliminary thoughts. In R. Steinberg (Ed.), *Music and the mind machine: The psychophysiology and psychopathology of the sense of music* (pp. 105-113). Berlin: Springer-Verlag.

Davidson, J. (1997). The social in musical performance. In D. H. North (Ed.), *The social psychology of music.* New York: Oxford Unviersity Press.

Davidson, J. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae scientae V* (2), 235-256.

Friberg, A. & Battel, G. U. (2002). Structural communication. In R. Parncutt, & G. E. McPherson (Eds.), *The Science and psychology of music performance: Creative strategies for teaching and learning* (pp. 199-218). New York: Oxford University Press.

Gabrielsson, A. & Juslin, P.N. (1996). Emotional expression in music performance: Between performer's intention and listener's experience. *Psychology of music, 24,* 68-91.

Gromko, J. E. (1993). Perceptual differences between expert and novice music listeners: A multidimensional scaling analysis. *Psychology of music, 21,* 34-47.

Hill, P. (2002). From score to sound. In J. Rink (Ed.), *Musical performance: A guide to understanding* (pp. 129-143). Cambridge, UK: Cambridge University Press.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and psychophysics, 14,* 201-211.

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

Juslin, P. & Persson, R. (2002). Emotional communication. In R. Parncutt, & G. E. McPherson (Eds.), *The Science and psychology of music performance: Creative strategies for teaching and learning* (pp. 219-236). New York: Oxford University Press.

Knapp, M. L. & Hall, J. A. (2002). *Nonverbal communication in human interaction* (5th edn.). USA: Wadsworth Thomson Learning.

Massaro, D. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. New Jersey: Lawrence Erlbaum Associates.

McClaren, C. A. (1988). The visual aspect of solo marimba performance. *Percussive notes, Fall,* 54-58.

Repp, B. (1999). Control of expressive and metronomic timing in pianists. *Journal of motor behavior* (electronic version). *31*.

ISBN 88-7395-155-4

Schaffer, L. H. (1995). Musical performance as interpretation. *Psychology of music, 23,* 17-38.

Sloboda, J. (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: The acquisition of expert performance in the arts and sciences, sports, and games.* New Jersey: Lawrence Erlbaum Associates.

Todd, N. P. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the acoustical society of America, 91,* pp. 3540-3550.

Vines, B.W., Krumhansl, C.L., Wanderley, M.M. & Levitan, D.J. (in press). Cross-modal interactions in the perception of musical performance. *Cognition*.

Wanderley, M., Vines, B., Middleton, N., McKay, C. & Hatch, W. (2005). The musical significance of Clarinetists' ancillary gestures: An exploration of the field. *Journal of new music research, 34* (1), pp. 97-113.

Wickens, C. D., Lee, J., Liu, Y. & Becker, S. G. (2004). *An introduction to human factors engineering* (2nd edn.). New Jersey: Pearson Prentice Hall

ISBN 88-7395-155-4