

AFFECTIVE AND COGNITIVE RESPONSES TO MUSICAL
PERFORMANCES OF EARLY 20TH CENTURY CLASSICAL
SOLO PIANO COMPOSITIONS: THE INFLUENCE OF MUSICAL
EXPERTISE AND AUDIO-VISUAL PERCEPTION

MARY C. BROUGHTON, & JESSIE DIMMICK
The University of Queensland, Brisbane, Australia

ROGER T. DEAN
Western Sydney University, Sydney, Australia

EFFECTIVE AUDIENCE ENGAGEMENT WITH MUSICAL performance involves social, cognitive and affective elements. We investigate the influence of observers' musical expertise and instrumental motor expertise on their affective and cognitive responses to complex and unfamiliar classical piano performances of works by Scriabin and Hanson presented in audio and audio-visual formats. Observers gave their felt affect (arousal and valence) and their action understanding responses continuously while observing the performances. Liking and familiarity were rated after each excerpt. As hypothesized: visual information enhanced observers' action understanding and liking ratings; observers with music training rated their action understanding, liking and familiarity higher than did nonmusicians; observers' felt affect did not vary according to their musical or motor expertise. Contrary to our hypotheses: visual information had only a slight effect on observers' arousal felt affect responses and none on valence; musicians' specific instrumental motor expertise did not influence action understanding responses. We also observed a significant negative relationship between action understanding and felt affect responses. Ideas of empathy in musical interactions motivated the research; the empathy framework in relation to musical performance is discussed. Nonmusician audiences might be sensitized to challenging musical performances through multimodal strategies to build the performer-observer connection and increase understanding of performance.

Received: April 10, 2019, accepted August 2, 2020.

Key words: musical performance and expertise, perception-action, audio-visual perception, felt affect, empathy

AUDIENCE MEMBERS REPORTEDLY DESIRE THE shared, communal experience that live performance offers (Dearn & Price, 2016; Radbourne, Johanson, Glow, & White, 2009). This is partly because the opportunity for social, and slightly challenging cognitive and emotional experiences motivates audience engagement and re-engagement with performing arts (Kemp & White, 2013; Radbourne et al., 2009; Tajtáková & Arias-Aranda, 2008; Walmsley, 2011). However, the audience member's degree of embodied expertise and knowledge about the art form can reportedly facilitate or hinder engagement (Dobson, 2010; Dobson & Pitts, 2011; Tajtáková & Arias-Aranda, 2008). Pitts (2005) highlights that the visual and social aspects of live musical performance are crucial elements that contribute to positive experiences for audience members. Furthermore, the personal social connection that the performer forms with the audience, such as through verbal introductions, can enhance the audience's responses to performance (reviewed in the context of both music and dance: Stevens, Dean, Vincs, & Schubert, 2014). On the other hand, the evidence that the provision of psychohistorical information per se is positive is currently unconvincing (Chmiel & Schubert, 2019).

This study complements the extant largely qualitative research on audience responses to classical musical performance with an experimental approach focusing on the musician-audience member (hereon termed "observer") connection during performance. The aim is to investigate observers' affective and cognitive responses in relation to the performer of unfamiliar and potentially challenging Western classical piano music compositions, and how these might vary depending on differences in observers' embodied musical expertise, and whether the performer can be seen and heard, or heard only. Throughout, it is important to bear in mind that the observer is likely always reacting to the musical sound, while their awareness of and responses to the performer may vary according to condition and disposition.

MUSICIANS' BODILY MOVEMENT AND PERFORMER-OBSERVER COMMUNICATION

Seeing and hearing a musician perform may heighten communication of both musical content and expression. While music is usually thought of as primarily an auditory experience, the visual component is highly powerful in performer-observer communication (Broughton & Stevens, 2009; Davidson, 1993, Vines, Krumhansl, Wanderley, & Levitin, 2006), just as gestural and nonverbal information is communicative and important in everyday social interactions (McNeill, 1992). Performing musicians' gestures can influence perception of note duration (Schutz & Lipscomb, 2007), expressive features such as phrasing, dynamics, and rubato (Juchniewicz, 2008), through to judgments of expressiveness and interest in performance (Broughton & Stevens, 2009). Additionally, Broughton and Stevens (2009) found that musically trained observers perceived musical performance excerpts to be significantly more expressive and interesting than did musically untrained observers. Evidently, the multi-modal performance experience benefits performer-observer communication, and observers' expertise affects their responses to musical performance. While the music is often the focus of performance, here we consider whether observers might actually be connecting with the performer, as the conduit for the music.

EMPATHY AND MUSICAL PERFORMANCE

A growing body of theory proposes that observers might respond to emotionally expressive musical experiences as they would empathize with another human (Miu & Vuoskoski, 2017). Broadly defined, empathy is an affective response to another that has some correspondence to the affective state of the other, that involves actual or inferred recognition and some experience of the other's affective state while a distinction between "self" and "other" is maintained (Decety & Jackson, 2004; Decety & Lamm, 2009; Fan, Duncan, de Greck, & Northoff, 2011; Ickes, 1997). Understanding is proposed to stem from the ability to perspective take, or project the "self" into others by putting oneself in another's shoes (Davis, 1983). Proposed key mechanisms involved in empathy include shared representations between actor and observer, underpinned by some common perception-action coding, and simulation or resonance mechanisms moderated by regulatory processing and self-other awareness (Decety & Jackson, 2004).

A recent model of musical empathic interactions (Wöllner, 2017), founded on the notion that musical interactions are social experiences, places the

perception-action circuit at its center. It proposes that the audience connection to performers involves social empathy developed through performer-audience/observer interactions, which can facilitate conscious perspective-taking. The model proposes that the music forms a second type of subject (along with performers) with which the audience can empathize. The music might be ascribed some type of abstract "persona" (Levinson, 2011) of which listeners attempt to take the perspective. A third element is necessarily the interaction between co-performers, performance agency, and the music. These ideas provide some motivation for the present study. However, operationalizing empathy in musical interactions, when the emotive topic is the music as much as the performer (object as much as person) is yet to be adequately defined in theory (see also the Discussion section). Responding to this, we operationalize specific features that might be components contributing to empathy, and consider their possible relationships to empathy in the Discussion. Thus the present study focusses on observers' affective response to musical performance and their cognitive understanding of a performing musician's expressive action.

AFFECTIVE RESPONSES TO MUSICALLY EXPRESSIVE PERFORMANCE

Prominent music and emotion theories propose that listeners respond to music by way of several distinct mechanisms that include bottom-up emotional contagion and top-down appraisal processes (see Juslin, 2013; Juslin & Västfjäll, 2008; Scherer & Coutinho, 2013). This suggests that listeners recognize and "mimic" affective expression from musical sound. Evidence from subjective behavioral, psychophysiological, and neural activation response studies indicates that individuals can share the identification and feeling of affects from listening to music that has been performed by humans, and particularly classical music (for a review, see Eerola & Vuoskoski, 2013). Furthermore, affective musical and vocal expressions appear to be communicated through similar patterns of acoustic cues (Cespedes-Guevara & Eerola, 2018; Juslin & Laukka, 2003). Of course, there are many potential mechanisms by which musical performance might induce an affective response in an observer. Amongst these, an observer's affective experience to a musical performance might well involve a connection with the musician, as the generator of musical sensory information.

Theoretical propositions such as the Shared Affective Motion Experience (SAME) model (Overy & Molnar-Szakacs, 2009) argue that observers/listeners affectively respond to music via a connection with the human-produced motion needed to create the musical sound.

This is proposed to involve a network of neurons in the temporal cortex, the fronto-parietal Mirror Neuron System and limbic system, which is similar to a network proposed to underpin empathy (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003). Shared activation across this neural network between music performer and observer/listener is also proposed to underpin affective “emotional contagion” (see Juslin, 2013; Juslin & Västfjäll, 2008) type responses to music. Furthermore, research indicates that we can recognize whole-body dynamic emotional expressions from patterns of characteristic motor elements (Shafir, Tsachor, & Welch, 2016), and observation of another’s emotional expressions can induce similar emotional states in observers (Shafir, Taylor, Atkinson, Langenecker, & Zubieta, 2013). It is beyond the scope of this study, if not impossible to completely separate observers’ affective responses to the performer from those due to the sound of the musical compositions. But the evidence suggests that where music is performed by humans there likely exists a shared performer-observer connection which plays a role in observers’ subjective felt affect responses.

Seeing as well as hearing a performing musician potentially enhances the relative contribution of the performer’s expressive intentions to observers’ felt affect relative to the composition and other affect-induction mechanisms. Indeed, viewing the musician performing has been claimed to have a powerful effect on observers’ felt affect, as measured through subjective measures, such as experienced tension (Vines et al., 2006), and physiological means (Chapados & Levitin, 2008). However, Vuoskoski, Gatti, Spence, and Clarke (2016) found that while observers’ skin conductance responses were greater when presented piano performance in an audio-only mode (as compared to audio-visual and visual-only), observers’ self-reported felt affect did not differ between audio-only and audio-visual presentation modes. Such contrasting results within previous research might be accounted for by the nature of the musical stimuli, which was a Romantic tonal musical composition (Vuoskoski et al., 2016) compared to an atonal composition (Chapados & Levitin, 2008), which more closely matches the more challenging musical style and era of the stimuli presented in the study reported here. Interestingly, observers’ musical expertise did not seem to influence their affective responses in these studies. Furthermore, music training appears to bear no influence on how individuals categorize affects induced through music listening (Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005). Therefore, while observers’ felt affect responses are not expected to vary on the basis of their musical expertise in this

study, the presence of visual as well as auditory information is expected to heighten perception of the performers’ embodied expression.

In the present study we attempt to control for potential influences of observers’ personal musical preferences on affective responses by presenting complex and unfamiliar music and measuring liking and familiarity with the stimulus material after each presentation. We attempt a balance of fast and slow-paced excerpts (Balch & Lewis, 1996; Husain, Thompson, & Schellenberg, 2002), and loud and quiet excerpts (Bailes & Dean, 2012; Dean, Bailes, & Schubert, 2011; Schubert, 2004) to control for potential effects of musical features on observers’ arousal responses. In addition, we use music that has no obvious mode (major/minor) that could potentially affect valence responses (Husain et al., 2002). However, individuals with music training might retrospectively rate the music more highly in liking and familiarity than untrained observers, and might cognitively connect with the performing musician more readily.

COGNITIVE UNDERSTANDING OF A PERFORMING MUSICIAN’S EXPRESSIVE ACTION

Witnessing another’s expressive action provides observers with the opportunity to perceive and understand, to some degree, the expressive state of the other person. Research suggests that shared embodied representations account for the communication of expressive goal-directed actions (Gallese, 2003). That is, shared experiences of actions, emotions, and sensations between people provide a neurobiological basis for interpersonal communication and understanding of others. Furthermore, when instructed, observers appear to be able to take the perspective of a performing musician so as to imagine how the performer feels in relation to the music they are playing; and this may influence the affective state experienced by observers (Miu & Baltes, 2012). However, the degree to which observers might be able to cognitively take the perspective of a performing musician, or “put themselves in their shoes” and understand their cognitive and affective state is likely shaped by the degree to which the observer and performer have shared embodied experiences and mental representations, hence the capacity for action understanding.

Training and embodied experience can shape neural representations for action production and perception, and perceptual and cognitive decision-making processes in different contexts. The results of experimental research using functional magnetic resonance imaging (fMRI) suggest that specialist motor training modifies human neural responses to artistic action stimuli, such as classical ballet or capoeira (a Brazilian martial

art-dance fusion; Calvo-Merino, Glaser, Grèzes, Passingham & Haggard, 2005). Specifically, the strongest activations are seen in observers' neural areas associated with production of familiar movements in line with their professional motor training (e.g., classical ballet or capoeira, Calvo-Merino et al., 2005). Expertise and training therefore create embodiments, or neural representations, that are not seen in untrained controls. The hypothesized human Mirror Neuron System (MNS) has been proposed as the mechanism involved in expertise-moderated action understanding, and also the mechanisms behind cognitive and affective components of empathy (Milston, Vanman, & Cunnington, 2013). According to this view, mirror neurons are active and fire when individuals observe or execute the same goal-directed action (Gazzola, Aziz-Zadeh, & Keysers, 2006; Milston et al., 2013). Observed actions are thus mapped onto equivalent representations in the observers' brains. Research suggests that expertise also shapes the way in which individuals attend and respond to multimodal cues in the environment. For example, pilots' expertise-derived mental models in long term memory appear to direct attention and moderate decisions for effective task performance (Beltenkes, Wickens, & Kramer, 1997; Doane, Sohn, & Jodlowski, 2004; Schriver, Morrow, Wickens, & Talleur, 2008). Therefore, the degree to which observers possess the embodied experience necessary to produce the action and carry out the task that they witness might be expected to shape their perception and understanding of the action.

Research in music indicates that observers' specialist music-related motor expertise may shape their patterns of neural activation, perception, and judgments of performing musicians' embodied expression. Using fMRI, Haslinger et al. (2005) found that professional pianist observers exhibited stronger neural activation in fronto-temporo-parietal regions than musically untrained controls in response to piano playing actions. Instrumental musical expertise also seems to affect attention and cognitive decision-making processing about musical performance. For example, expert musicians independently applied an analytical system (Laban effort-shape analysis), following training, to analyze the embodied expression they perceived in audio-visual recordings of solo marimba¹ performance

¹The marimba is a wooden keyboard percussion instrument. The keyboard layout is similar to a xylophone, but it has a deeper and wider pitch range. It spans a five-octave range and measures approximately two-and-a-half meters in length. Solo marimba players perform piano-like music with one or two mallets in each hand.

(Broughton & Davidson, 2014; Broughton & Stevens, 2012). Results suggested that observers with differing instrumental motor expertise noted many expressive moments at similar locations in the performance material. However, their analysis and categorization of the performers' embodied expression at these expressive moments differed according to their experience in marimba playing. This suggests that the ability to cognitively take the perspective of a performing musician and understand their goal-directed expressive action might be enhanced where the observer shares the same music-related motor expertise with the performer (particularly when they play the same instrument).

RESEARCH AIM, DESIGN, AND HYPOTHESES

This study aims to investigate how observers' felt affect and action understanding responses to the performance of early 20th century Western classical solo piano compositions differ according to observers' musical and specific motor expertise, and the modality of presentation. The computer-based experiment is a 3 (Expertise: musician-pianist, musician non-pianist, nonmusician) x 2 (Modality of presentation: audio-only, audio-visual) mixed between-within repeated measures design. No differences for felt affect (arousal, valence) are expected between expertise groups; audio-visual modality of presentation is expected to enhance observers' felt affect responses. It is expected that musician pianists will report higher action understanding ratings than musician non-pianists who will, in turn, report higher action understanding ratings than nonmusician observers. Action understanding and liking ratings are expected to be higher for audio-visual presentations in comparison to audio-only. Musician pianists and musician non-pianists are expected to report higher liking and familiarity ratings than nonmusician observers. A relationship between action understanding and arousal-valence measures is expected, as cognitive and affective processes are co-active in normal situations, which is enhanced by the audio-visual condition.

Method

PARTICIPANTS

A total of 75 observers (age range = 17–51 years; 20 males, $M_{\text{age}} = 23.65$ years, $SD = 9.13$; 55 females, $M_{\text{age}} = 20.18$ years, $SD = 3.46$) voluntarily participated in the experiment. Scores on the Interpersonal Reactivity Index (IRI) (Davis, 1983) and Autism Spectrum Quotient Short Form (AQ-10) (Baron-Cohen, Wheelwright,

Skinner, Martin, & Clubley, 2001) were included to screen for sound interpersonal empathic competence. All observers fell within the normal range on both of these tests.

As in previous research (Wöllner & Cañal-Bruland, 2010), observers were grouped according to their musical and instrumental expertise. Observers' self-identification in one of three expertise groups—musician-pianist, musician (non-pianist), nonmusician—was verified by questionnaire. Each expertise group consisted of 25 observers. Previous research has demonstrated that expert performance at an international level in any field requires approximately 10,000 hours of sustained and deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993; Krampe & Ericsson, 1996). This is equivalent to almost three hours of practice a day for 10 years. The current study, however, employed a less stringent classification of expertise, considering that the observers were predominantly undergraduate university students.

Observers that self-identified as possessing more than seven years of formal music training were classified as musically trained. Observers who did not meet this threshold were still included in the musically trained groups if they fulfilled one or more of the following criteria: attainment of a Grade 7 or higher Australian Music Examinations Board (AMEB) practical examination on their primary instrument, or an Associate of Trinity College London (ATCL) performance diploma as this standard is considered entrance-level for undergraduate music degrees. Observers who reported being currently active in playing or performing their instrument on a regular basis (i.e., several times a week) and self-identified as performing, teaching, or composing musicians were also included (see Zhang & Schubert, 2019).

Musician-pianist. Musician-pianist ($n = 25$) observers self-reported that their principal instrument was piano, then had completed a minimum of seven years of formal piano training ($M = 13.22$ years, $SD = 7.30$, range = 7–40 years), and were currently active performing, teaching, or composing music. Three observers did not report their years of formal piano training but self-reported piano as their primary instrument, self-identified as performing musicians, and had attained a high level of AMEB practical examinations (Grade 6 and 8). The participant who reported Grade 6 AMEB was still included on account of self-reported formal piano training throughout primary and secondary schooling years.

Musician (non-pianist). Musician (non-pianist) observers ($n = 25$) self-identified as musicians whose primary instruments were not piano. They self-reported several years of formal instrumental music training on a primary instrument other than piano (M

= 9.04 years, $SD = 2.81$, range = 6–12 years). Two observers reporting possessing only six years of formal instrumental training were still included in this group on account of their self-identification as musicians, attainment of Grade 7 AMEB practical examination, and regular performance activity. Twenty musician (non-pianists) reported five years or fewer of formal piano training ($M = 0.75$, $SD = 1.71$, range = 0–5 years). Five observers who reported more than five years of piano playing experience (range = 9–17 years) were included in the non-pianist group because they self-identified as musicians whose principal instruments were not piano, nor did they play piano with any regularity. To illustrate, they were currently training at a tertiary level (e.g., Bachelor of Music, Queensland Conservatorium) and majoring in primary instruments other than piano.

Nonmusician. Twenty-five musically untrained observers self-identified as nonmusicians and had undertaken less than two years of formal music training ($M = 0.8$ years, $SD = 0.99$).

Observers' music preferences. Observers self-reported their music preferences via the Short Test of Musical Preferences – Revised (STOMP-R) (Rentfrow & Gosling, 2003). The STOMP-R assesses liking of a variety of music genres organized under four different dimensions: reflective and complex, energetic and rhythmic, upbeat and conventional, and intense and rebellious. The frequency of liking response for each genre and dimension according to each expertise group are summarized in Table 1. Musician-pianists (43%), more so than both the musician (non-pianist) (25%) and nonmusician observers (19%), preferred the *reflective and complex* dimension that encompasses the classical genre. Interestingly, musicians (non-pianists) differed from musician-pianists, preferring both the *energetic and rhythmic* and *intense and rebellious* dimensions more than genres in the *reflective and complex* dimension. Nonmusicians mostly preferred genres in the *energetic and rhythmic* preference dimension.

Observers were recruited through musician networks, a student research sign-up online system, and through print advertisements on campus. Observers received \$10 reimbursement for their time and travel expenses associated with participating in the research, or course credit. Observers with self-reported normal or corrected-to-normal vision and normal hearing were included in the study.

STIMULI

The stimulus material was drawn from a live recording of a concert given by a renowned Australian pianist and contemporary music specialist at The University of

TABLE 1. Observers' Self-Reported Music Preferences Gathered using the Short Test of Musical Preferences – Revised (STOMP-R, Rentfrow & Gosling, 2003)

Genre	Musician-pianist (<i>n</i> = 25)	Musician (non-pianist) (<i>n</i> = 25)	Nonmusician (<i>n</i> = 25)
Reflective & Complex			
Bluegrass	4	0	1
Blues	11	12	5
Classical	21	12	9
International/Foreign	11	3	7
Jazz	19	14	5
New Age	2	3	5
Opera	11	4	1
Folk	7	2	5
Total likes (% of each expertise group)*	86 (43%)	50 (25%)	38 (19%)
Energetic & Rhythmic			
Funk	1	11	10
Dance/Electronica	8	13	13
Rap/Hip-Hop	4	13	14
Reggae	5	11	4
Soul/R&B	10	11	12
Total likes (% of each expertise group)	28 (22%)	59 (47%)	53 (42%)
Upbeat & Conventional			
Religious	4	1	2
Gospel	5	3	3
Country	6	2	5
Oldies	6	11	8
Pop	13	17	15
Soundtracks/Theme Songs	18	19	16
Total likes (% of each expertise group)	52 (35%)	53 (35%)	49 (33%)
Intense & Rebellious			
Punk	3	9	4
Heavy Metal	3	6	4
Alternative	6	16	15
Rock	8	16	11
Total likes (% of each expertise group)	20 (20%)	47 (47%)	34 (34%)

Note: Summary of frequency of genre liking by each music preference dimension and expertise group for scores on the STOMP-R.

*(X% of each expertise group) refers to the proportion of participants from each expertise group that identified a preference for one or more genres in each dimension. Participants were able to identify as many genres as they preferred in the list provided. However, they were only counted once in determining the proportion of participants from their expertise group that reported a preference for music in the particular dimension. In contrast, the "total likes" expressed is simply the sum of all the rows above, in all but one case exceeding the number of participants in the specified expertise group.

Queensland (UQ) School of Music. Eight excerpts of music from three early 20th century classical music pieces were selected. The pieces were Sonata No. 9, Op. 68 "Black Mass" (1913) and Poème "Vers la Flamme" Op. 72 (1914) by Russian composer Alexander Scriabin, as well as Sonata (1940) by Australian composer Raymond Hanson. All pieces demanded a high degree of proficiency to perform. Excerpts were recorded in an audio-visual (AV) format, taking into view the length of the piano and full height of the seated performer from the side.

The audio-visual recording was edited to make a total of eight, 56–60 second selections (excerpts) that

included mostly complete musical phrases. An effort was made to select excerpts that reflected a balance of musical elements in order to elicit a range of affective responses: tempo (fast, slow), range of movement (large, constrained), and dynamics (loud, quiet) (see Davidson & Edgar, 2003; Schellekens & Goldie, 2011). Each of the AV computer files (.avi) were then also converted into audio files (.wav). Each of the eight excerpts were presented in two sets. In one set, the eight excerpts were presented twice audio-visually, and in the other set twice as audio-only (i.e., observers saw a black screen while the sound played).

APPARATUS AND MATERIALS

Excerpts were performed on a Steinway grand piano. Recordings were made on a Sony HDR-XR550 digital video camera featuring the Audio Video Coding High Definition (AVCHD) recording format for high definition video and stereo audio with 48 kHz sampling. Video editing and conversion of AV computer files (.avi) into audio (.wav) files was performed using Adobe Premier Pro CC 2014. Presentation[®] software was used to present the experiment and gather observers' responses. Stimuli were displayed on a Dell U2414H monitor running at 60 Hz. Audio was provided through Bose (QuietComfort[®] 25 Acoustic Noise Cancelling[®]) headphones at a comfortable listening level. Continuous self-report judgements were made using a Logitech Attack 3 Joystick (J-UG18) with USB 2.0 connector, ambidextrous handle, responsive control, and lower spring force, which applies a small degree of resistance as the joystick is maneuvered away from the neutral, upright position. This assists participants to know where the neutral position of the response scale is by feel.

Demographic and music background questionnaire. Observers' demographic (i.e., age and gender) and musical background information (e.g., formal music and instrumental training) was collected by a questionnaire designed for the study presented using Qualtrics online survey software.

IRI and AQ-10 questionnaires. The IRI (Davis, 1983) measures affective and cognitive components of empathy. It consists of four subscales (perspective taking, empathic concern, fantasy, personal distress) each with seven items. Each subscale demonstrates high internal reliability indices of $\alpha = .70$ to $.78$ (Davis, 1983). For males, the correlation between the test and retest scores ranges from $r_s = .61$ to $.79$, and from $r_s = .62$ to $.81$ for females (Davis, 1983). There were fewer than 20% missing values, so the mean of the subscales were taken to substitute for the missing responses (see Hills, 2003).

The AQ-10 (Baron-Cohen et al., 2001) is often used as a quick referral guide for adults who do not have a learning disability but are suspected of having an autism spectrum disorder. It was used as a screening measure because empathy and social-affective interpersonal competency is believed to be impaired in those with autism (e.g., Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007). The scale has demonstrated good internal consistency ($\alpha = .72$) (Sizoo et al., 2015). Observers responded to the 10 items on a 4-point Likert scale, ranging from 1 (*definitely agree*) to 4 (*definitely disagree*). Individuals who scored greater than six were excluded from analyses.

Felt affect measures. Observers indicated *felt emotion* (being their felt response and not the emotion expressed by the music) continuously along two concomitant dimensions (*arousal* and *valence*) as the performance excerpt unfolded by moving a joystick in two-dimensional space. The word "emotion" was presented to participants, rather than "affect," as it is more readily comprehensible by a person inexperienced in psychological experiments. We primarily use the word "affect" in the text because it is less prescriptive, and avoids unintentional connotations the term emotion might have in different areas of psychology. Felt affect is elicited in a music observer/listener, as distinct from that expressed by a musician. The two-dimension arousal (ranging from calming to arousing) valence (from positive to negative) emotion space is a supported (see Russell, 1980), reliable, and well-used means of measuring continuous affective responses to music (Nagel, Kopiez, Grewe, & Altenmüller, 2007). Previous research asking observers to respond continuously on one dimension to musical performance stimuli reports that the task does not interfere greatly with observers' responses (Egermann, Pearce, Wiggins, & McAdams, 2013; Stevens, Vincs, & Schubert, 2009), and listeners in several studies have successfully made continuous self-report arousal and valence responses simultaneously to musical performance stimuli (e.g., Bailes & Dean, 2012; Grewe, Nagel, Kopiez, & Altenmüller, 2007; Schubert, 1999, 2004).

Action understanding measure. Action understanding in the context of this study is defined as the degree to which observers felt that they could put themselves in the performer's shoes and understand what the performer was doing to physically generate the expressive performance. The decision to use of the term "expressive" was based on the notion that it is usual to refer to "expressive musical performance," more so than "emotional" or "affective" musical performance. Additionally, understanding of the word "emotional," as used in common parlance, might conjure positive or negative sentiments and potentially confound results, which we wanted to avoid. Action understanding was self-reported continuously as each performance excerpt was presented. The action understanding measure expected observers to take the perspective of the performer, in the nature of the cognitive facet of empathy. Observers were asked to maneuver the joystick left and right, ranging from *do not understand at all* to *understand completely* to reflect their rating.

Liking and familiarity measures. Liking of and familiarity with the music just heard was reported for each excerpt at their completion on separate 7-point Likert

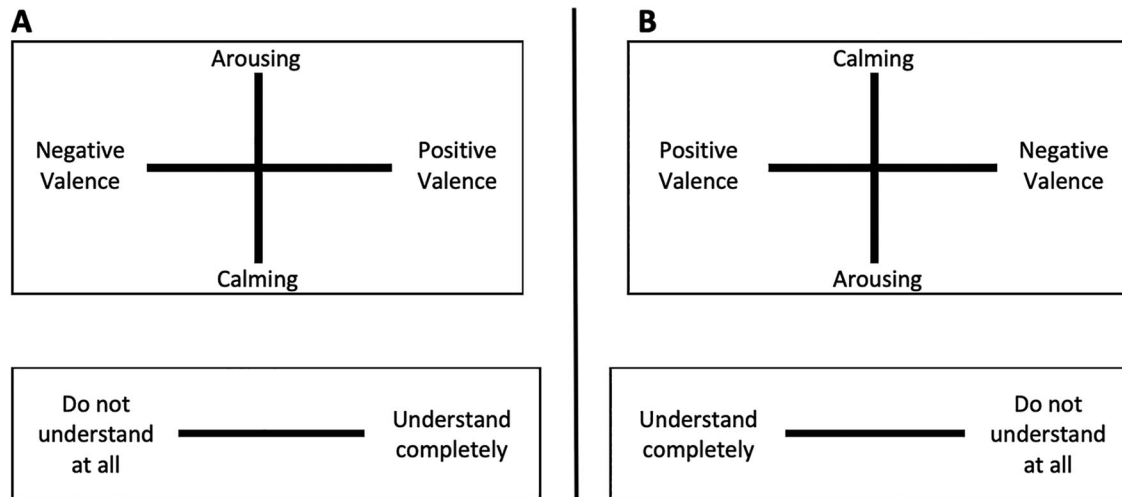


FIGURE 1. Counterbalancing of felt affect (arousal and valence) and action understanding response axes. Panel A shows the original arrangement of axis labels. Panel B shows the inverted axis label arrangement.

scales, ranging from 1 (*dislike very much*) to 7 (*like very much*) and from 1 (*completely unfamiliar*) to 7 (*very familiar*), respectively. Liking and familiarity were primarily measured to gain insight into observers' preferences as an indicator of experience with and long term memory for a similar style music. In addition, although the music was novel initially, Zajonc (2001) has documented that repeated exposure to a stimulus can prompt increases in positive affect and preference. Because felt affect is posited to be involved in music preferences (Schubert, 2007), the repetition of each excerpt in the present study could potentially moderate both liking and familiarity responses. Therefore, *liking* as well as *familiarity* were also measured to help account for any effects of exposure on affect, should differences between expertise groups be observed.

PROCEDURE

Observers gave written informed consent to participate in the research. Ethical approval to conduct the research was obtained from The University of Queensland Behavioural and Social Sciences Ethical Review Committee. To control for any stimulus order effects, the excerpts were randomized within each condition for each observer. In addition, the order of responding and modality of presentation were counterbalanced for each observer. Counterbalancing was also employed for the axes for both felt affect and action understanding responses. This was done to avoid any moderating effects of approach and avoidance movements for the

joystick response on felt affect responses. That is, previous research has shown that making approach and avoidance movements relative to the self are analogous to the "pursuit of pleasure and avoidance of pain" (Elliot, 2006, p. 111). Observers' were randomly assigned one of the possible arrangements of the axes for both felt affect and action understanding responses for each condition. The labels on each response axis were inverted to counterbalance joystick response movements (see Figure 1).

Observers completed the study on an individual basis in a quiet office space. Upon arrival, observers were seated at a computer and completed the informed consent procedures prior to commencing the research tasks. Observers then completed the music and demographic study questionnaire, as well as the STOMP-R, IRI, and AQ-10.

For the experimental task, the stimulus material was presented on the computer monitor, with the audio delivered through headphones at a comfortable listening level. Two samples were recorded per second, which is standard for continuous self-report judgements (Schubert, 2006). The two sets, one audio-visual (AV) and one audio-only (AO), were counterbalanced across observers, with eight excerpts randomized within each. All excerpts were presented twice within each set: continuous self-report judgements of felt affect (hence, arousal and valence) were made on one presentation, and action understanding on the other. The order of these responses was also counterbalanced. The AV and

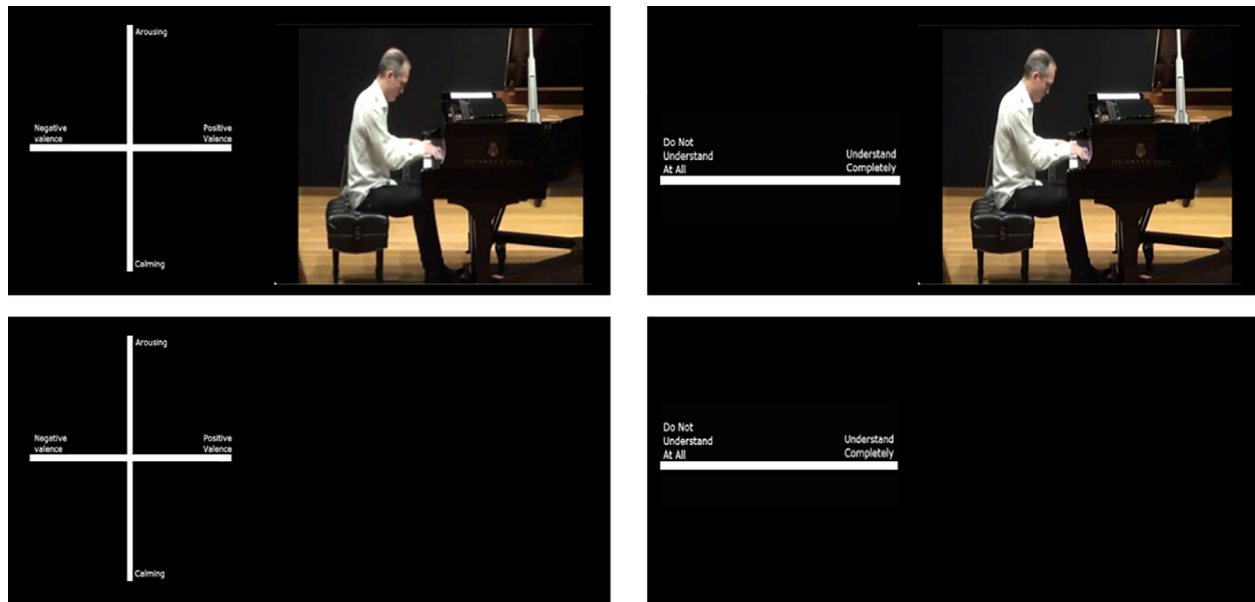


FIGURE 2. AV and AO conditions stimulus presentations for felt emotion and action understanding responses.

AO stimulus presentations are illustrated in Figure 2. At the completion of each excerpt, the experimental program prompted the observers to make their liking and familiarity judgements on two separate Likert-scales by pressing a number key on the keyboard between one and seven that best fit their response in each case. This process was repeated until each of the eight excerpts had been presented and responded to four times, twice AV and twice AO.

The measures (arousal and valence to indicate felt affect and action understanding) were explained to participants. Participants were instructed how to make arousal and valence responses by maneuvering the joystick around four quadrants of spatial movement in the two-dimension emotion space, analogous to four, quarter segments of a clock face. Exactly where the joystick was maneuvered within each segment would provide nuance as to the degree of arousal and valence felt. Participants were also instructed how to move the joystick between the far left and far right extremities of range to indicate their action understanding responses. The joystick's vertical resting position indicated the midpoint on the action understanding scale, or the midpoint intersection of the arousal and valence scales. Observers completed two training trials before beginning the experimental phase in order to familiarize themselves with the procedure and joystick response. Following training, observers were encouraged to ask questions or gain clarification regarding the procedure

before they commenced the main experiment. During the experiment, observers were able to pause the program if they needed a break, in addition to the scheduled pause that separated the AV and AO sets. Each session lasted for approximately one hour.

DATA PREPARATION

Observers' continuous response data from the experiment were logged electronically using the joystick on scales ranging from -250 to $+250$. Action understanding was logged according to joystick movement on the x -axis. Arousal was logged through joystick movement on the y -axis, and valence through movement on the x -axis according to the two-dimension emotion space (Russell, 1980; Schubert, 2006). Prior to analyses, however, all of the log data was shifted by adding 250 to each data point to put zero at the origin (i.e., 0 to $+500$). This generated separate time-series data on scales ranging from 0 – 500 for arousal, valence, and action understanding dependent variables.

DATA ANALYSIS

Multi-level mixed effects analyses were undertaken in R using the lme4 package; for liking and familiarity, these were analyses of point data; for the time series data, these constituted so called "cross-sectional time series analysis," abbreviated CSTSA. CSTSA is a mixed-effects method for simultaneous analysis of multiple time series, which does not require any data

averaging; the integrity of every individual data series is maintained. Mixed-effects analyses allow the separation of “fixed” effects (which represent participant population features) and “random effects” (which represent variation between individuals or pieces). The inclusion of random effects enhances the power of the analysis, the statistical strength of analysis of fixed effects, and also allows the direct study of interindividual variation when required.

In time series analysis, the fundamental concern is that sequential actions (i.e., individual successive data points) are highly autocorrelated: which means that any data point of such a time series is partially predicted by some combination of its prior data points, and hence data points are not statistically independent. Conversely, conventional statistical approaches assume that all data points are independent. Thus conventional statistics cannot be applied meaningfully (cf. Dean & Dunsmuir, 2016; Yule, 1926). For most types of time series analysis (TSA) it is necessary to obtain data series that are statistically “stationary”: that is, essentially, that show constant variance and constant covariance between data at different time points. Often (as described below), initial data series are non-stationary, and stationarity is achieved by differencing: which involves constructing a new series (one item shorter) as the differences between successive pairs of values of the original series. A huge body of literature has defined methods for time series analysis, and Dean and Bailes (2010a) have provided a detailed introduction to its use in the analysis of continuous responses to music. Mixed effects CSTSA enables the researcher to analyze autoregression (assessing how preceding values in a timeseries predict the present value of the time series), as well as fixed and random effects to arrive at the best model for the time-series data.

Restricted maximum likelihood (REML) was used to fit the models. The quality of each model was assessed in terms of goodness of fit and by the Bayesian Information Criterion (BIC). Goodness of fit is indicated primarily by the magnitude of the residual error (that portion of the data values that are not correctly modeled) which decreases as model fit improves. The BIC on the other hand is an estimate of the efficiency of a model, which penalizes models not only for poor fit, but also for the number of predictors it requires. In all cases the quality of the model residuals was acceptable as judged by Q-Q plots and assessment of normality. That is, the observed quartile points in the Q-Q plots did not clearly deviate from normality. In the time series cases, the lack of significant partial autocorrelation in the residuals also supported the models’ goodness-of-fit

to the observed data and indicated no need for further model complexity.

Results

DESCRIPTIVE STATISTICS

Figure 3 shows notched-boxplot summaries of the action understanding, liking, familiarity, arousal, and valence responses by expertise and modality of presentation. For the continuous response measures, for each individual response series a single mean value is obtained to enter into the summary dataset; these values are also used in the next subsection, *Mixed Effects Modeling*. When the notches of a pair of plots do not overlap, this indicates a significant difference in their medians. As expected, action understanding, liking and familiarity are all clearly higher in the musician-pianist (MP) and musician (non-pianist) (M) than in the nonmusician (NM) group. Contrary to expectations, the familiarity responses seem to be graduated across all three groups (MP > M > NM). Also contrary to expectations, familiarity responses seem to be enhanced by the AV condition in comparison with the AO. In contrast, but again in agreement with expectations, there are no obvious impacts of expertise on perceptions of arousal and valence: these thus probably reflect population tendencies rather than being distinguished amongst the expertise groups. Contrary to expectations, the AV condition does not enhance arousal and valence responses. Liking, familiarity and action understanding ratings were significantly correlated with each other (Spearman correlations .38 for liking and action understanding, .19 for familiarity and action understanding, and .29 for familiarity and liking; all $p < .001$), confirming their mutual relationships. This may explain why the familiarity responses are somewhat contrary to our expectations.

A comparable analysis was made of the coefficients of variation (abbreviated CV, which is measured as $SD/mean$) of the time series data (arousal, valence, and action understanding). CV is a simple normalized measure of variability in a data set (whether point data or time series data), and so it was used to consider whether expertise allows more nuanced (widely varying) responses as possibly indicated by larger CVs. The only significant difference observed was that in the AO condition the NM group showed higher CV for action understanding than did the M and MP groups; this was no longer true in the AV condition, again suggesting that the NM group had considerably more difficulty in action understanding in the absence of AV cues than did the two musician groups.

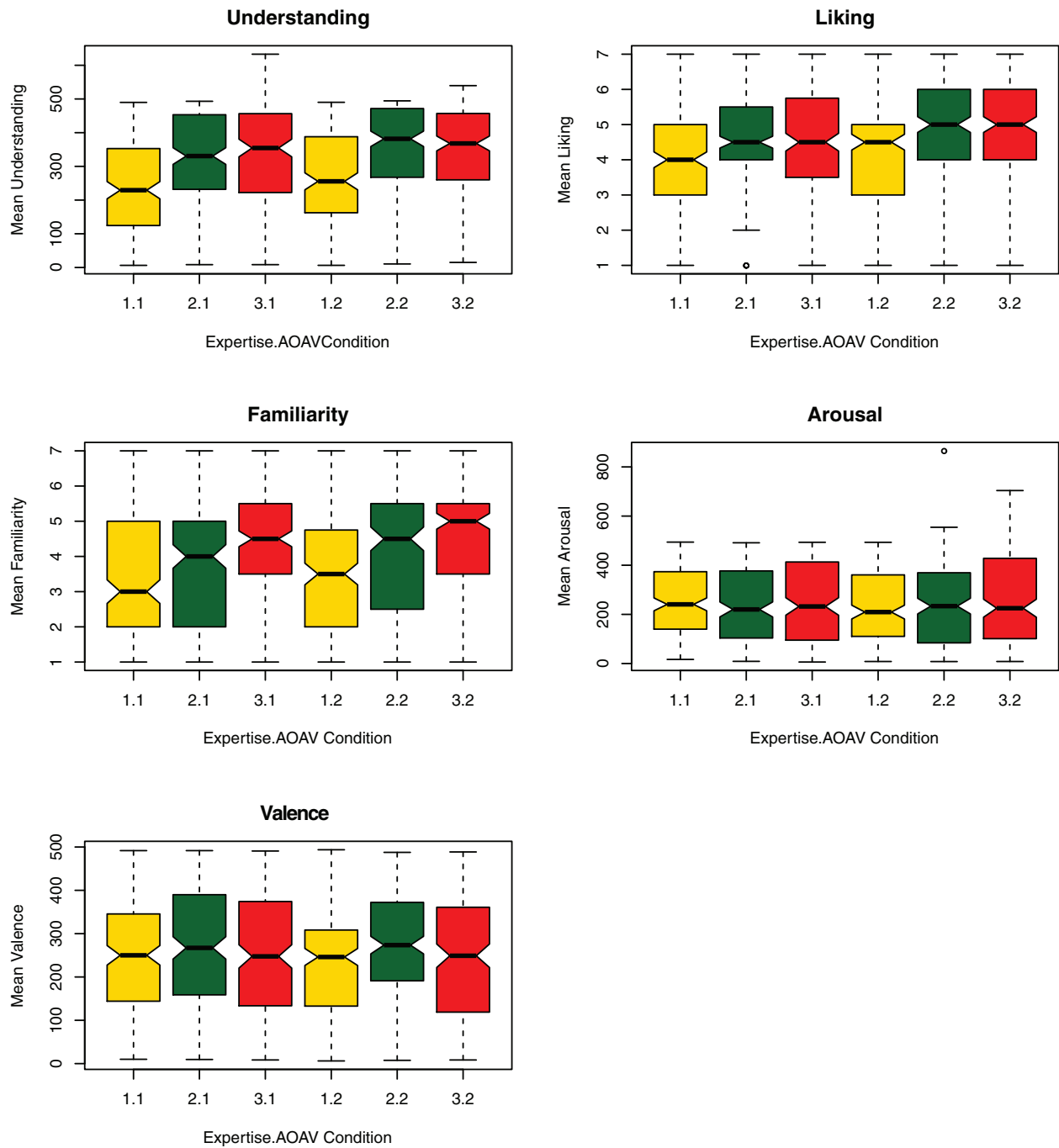


FIGURE 3. Boxplots of mean Action understanding (Understanding), Liking, Familiarity, Arousal and Valence Ratings by Expertise group and AO/AV condition. On the x-axis, the number preceding the decimal point refers to the three expertise groups: 1 = Nonmusician (NM), 2 = Musician (M), 3 = Musician-Pianist (MP). The number following the decimal point on the x-axis refers to the modality of presentation: 1= Audio-Only (AO), and 2 = Audio-Visual (AV). For example, "1.1" refers to Nonmusician.Audio-Only condition. Notches indicate plot medians.

TABLE 2. Fixed Effects for the Three Separate Mixed Effects Models for Action Understanding, Liking, and Familiarity

	Fixed effects in model	Value/Coefficient	SE	t	Correlation between model and data
Action understanding	Intercept	233.60	23.83	9.80	.78
	Musician	85.63	26.18	3.27	
	Musician-pianist	84.77	26.18	3.24	
	AV vs. AO	33.13	5.59	5.93	
Liking	Intercept	3.84	0.25	15.24	.74
	Musician	0.75	0.24	3.12	
	Musician-pianist	0.65	0.24	2.72	
	AV vs. AO	0.31	0.06	5.51	
Familiarity	Intercept	3.57	0.24	14.61	.74
	Musician	0.43	0.34	1.27	
	Musician-pianist	0.87	0.34	2.56	
	AV vs. AO	na	na	na	

Note: For each model, the (absolute) Values or Coefficients that are more than twice as large as the associated SE (ratio shown in the *t* column) are conservatively read as statistically significant at $p < .05$ level. Some of the coefficients are significant at lower probability levels, but the critical factor is that the coefficients (i.e., effect sizes) of the statistically significant predictors (all but one in the table) can be considered in relation to the scales on which the modeled values are expressed (e.g., the Likert ranges). The two musician group coefficients are with reference to the nonmusician group. Random effects (intercepts) were included in each model for participant and piece (not shown). na = not included in model.

MIXED EFFECTS MODELING OF AROUSAL, VALENCE, ACTION UNDERSTANDING, LIKING AND FAMILIARITY

Table 2 summarizes the significant results of the mixed effects models made with the lme4 library in R, specifically directed at testing our hypotheses concerning the influence of expertise (MP/M/NM) and modality of presentation (AO/AV). Note again that both fixed (population) and random (individual participant or item intercepts) effects are included in the model as described above in the *Method* section. As is common in mixed effects models, the random effects account for the majority of the explained variance, and this strengthens the interpretation of the fixed effects (the factors of interest here). Correspondingly, “null” models with only the random effects show quite good fit, and correlations with data similar to those shown in Table 2. More importantly, linear mixed models (without random effects, using the R package lm, since lme4 will not run without random effects), show quite similar coefficients and significances for the fixed effects to those in the mixed models illustrated, though the correlation of model:data are reduced to between 0.26 and 0.36. Neither expertise nor AO/AV condition were significant predictors in models of mean arousal or valence, consistent with the indications of Figure 3. The model for action understanding confirms the strong positive influence of expertise suggested by Figure 3, and indicates that the coefficients for the MP and M group compared with the base NM group are very similar. The AV condition also had a strong positive influence compared with the AO. The correlation between model and data was .78. As suggested by Figure 3, the model for liking

shows quite parallel effects to that for action understanding (correlation between model and data = .74). The model for familiarity again showed an effect of expertise, mainly driven by the MP group, but contrary to the impression from Figure 3, there was actually no significant effect of AO/AV condition, which is readily comprehensible and in accord with our expectations. There were no significant interaction effects between the two IVs, Expertise and AO/AV, in any of the models.

These models, like most in the literature, disregard the fact that Likert ratings are ordinal or more likely monotonic (i.e., not necessarily uniformly spaced), rather than continuous. However, a Bayesian monotonic regression with mixed effects (done in R using the package “brms”) was strongly confirmatory for the liking model.

CROSS-SECTIONAL TIME SERIES ANALYSIS OF AROUSAL AND VALENCE IN RELATION TO ACTION UNDERSTANDING

Although we observed that expertise and AO/AV conditions do not influence mean levels of arousal and valence, theories of empathy would suggest that the level of action understanding evinced by an observer might be a positive influence on their affective responses. Vector autoregression (VAR), a multivariate form of time series analysis, can be used to assess bidirectional interactions between dependent variables, so called endogenous variables in VAR, but software to do this with mixed effects is limited. Instead, cross-sectional time series analysis (CSTSA), with the assumption of linear responses and models, allows an assessment of the suggested influence of action understanding upon felt

TABLE 3. Fixed Effects for the Two Separate Mixed Effects Models based on Cross-Sectional Time Series Analysis for dArousal and dValence

	Fixed effects in model	Coefficient	SE	<i>t</i>	Correlation between model and data
dArousal	Time	3.593e-05	6.221e-06	5.78	.35
	AOAV1	-8.765e-01	4.675e-01	1.88	
	AOAV2	-1.099e+00	4.675e-01	2.35	
	dUnderstanding	-1.341e-02	2.330e-03	5.76	
	Lag 1 dUnderstanding	-9.089e-03	2.409e-03	3.77	
	Lag 1 dArousal	3.378e-02	4.273e-02	0.79	
	Lag 2 dUnderstanding	-1.178e-02	2.338e-03	5.034	
	Lag 2 dArousal	-1.314e-01	2.725e-03	48.21	
	Lag 3 dArousal	-5.269e-02	2.652e-03	19.87	
	Lag 4 dArousal	-3.488e-02	2.621e-03	13.31	
dValence	dUnderstanding	-0.008486	0.002192	3.87	.40
	Lag 2 dUnderstanding	-0.011708	0.002286	5.12	
	Lag 2 dValence	-0.087825	0.008661	10.14	
	Lag 3 dUnderstanding	-0.005418	0.002373	2.28	
	Lag 3 dValence	-0.088710	0.002776	31.96	
	Lag 4 dUnderstanding	-0.003972	0.002307	1.72	
	Lag 4 dValence	-0.044222	0.002603	16.99	

Note. For each model, the (absolute) Values or Coefficients that are more than twice as large as the associated SE (ratio shown in the *t* column) are again conservatively read as statistically significant at $p < .05$ level, and higher *t* values attain low *p* values. All but two of the shown predictors are significant, and those two are either required comparison levels or provided substantial benefit to the overall model (such that its quality was worsened by removal). In these time series data, the number of data points is relatively large, strengthening these interpretations. Then most importantly, the coefficients (i.e., effect sizes) of the statistically significant predictors can be considered in relation to the scales on which the modeled values are expressed (e.g., the Likert ranges). Random effects (not shown) were included in the dArousal model for piece (intercepts) and by-participant random slopes for the effect of Lag 1 dArousal. By-participant random slopes for the effect of Lag 1 dValence, Lag 2 dValence were included in the dValence model. AOAV1 = audio-only condition; AOAV2 = audio-visual condition. Besides autoregression, the lagged fixed effects represent the influence of the endogenous (self-reported) variable time series on the modeled dArousal or dValence time series with a delay of 1-4 samples (lags) between the two time series. Smaller lags reflect a closer temporal alignment between two time series (lag 0 being a perfect temporal alignment, indicated for example as dUnderstanding). In autoregression, or lagging one time series against itself (e.g., dValence and lag 1 dValence), often the predictive effect of the lagged time series on the present value time series decreases as lags increase, but since each coefficient ultimately impacts on almost entirely the same sequence of values (bar the omitted lags), a rough impression of the overall effect of a predictor, such as dArousal, can be obtained by summing the coefficients of its lags. A positive/negative coefficient suggests that the particular fixed effect increases/decreases dArousal or dValence.

arousal and valence. The response data were not statistically stationary (that is, they did not show the required constant variance and covariances between data at each given time lag), and so the models were made on the first differenced (stationarized) variables. The difference version of series "Test" is labelled "dTest." The CSTSA models (selected by the procedure described in the *Method* section) for the dArousal and dValence time-series data included the following autoregression, fixed, and random effects. The autoregression component of the model for dArousal used lagged dArousal time series data (i.e., the dArousal time series with successive delays of 1-4 samples to create lag 1 . . . lag 4 dArousal time series). For the dValence model, lagged dValence time series data (i.e., lag 2 . . . lag 4 dValence time series) was used to predict the dValence time series. Included in the dArousal model were fixed effects for Time (as the music and participants' responses unfold over time), dUnderstanding and lag 1 . . . lag 2 dUnderstanding time series, audio-only and audio-visual conditions. Random effects for piece, lag 1 dArousal, and participant were also included. In the dValence model, fixed

effects of dUnderstanding, and lag 2 . . . lag 4 dUnderstanding were included. Random effects included were lag 1 . . . lag 2 dValence, and participant. Table 3 summarizes the results of CSTSA models of change in arousal (dArousal) and change in valence (dValence).

The model for dArousal shows strong autoregression and negative coefficients on dUnderstanding and its first lag, suggesting that increases in action understanding create decreases in arousal. Arousal seemed to increase slightly with time for each piece, and only AO/AV condition 2 (audio-visual) was individually significant. This has to be considered in the context of the evidence above that AO/AV condition influences action understanding itself, and this cannot be analyzed in CSTSA: thus, no deduction can be safely made from it. Expertise was not required in this CSTSA model. The correlation between the model fit and the dArousal data was .35, suggesting that only about 12.5% of the variance of dArousal was explained, which is not surprising given the lack of consideration of the possible interactions with the other perceptual dependent (endogenous) variables. There was a random effect

on piece, indicating that unsurprisingly they differ notably in their relation to arousal. A similar effect of dUnderstanding on dValence is shown in the second model, with about 15.6% of variance explained ($r = .40$). AO/AV condition, piece, and expertise were not significant here. Interactions between the manipulated IVs, expertise, and AO/AV were not significant in either model.

Discussion

This study investigated how observers' affective and cognitive responses to the solo performance of early 20th century Western classical piano compositions, conceptualized as felt affect and action understanding, vary according to their musical and motor expertise. In addition, we investigated how being able to see and hear the performing musician might influence observers' responses in comparison to hearing only. As hypothesized, observers' felt affect responses did not vary according to their musical or motor expertise. We hypothesized and found that musically trained observers gave higher action understanding, liking, and familiarity responses than nonmusicians. However, contrary to our prediction, musicians' specific instrumental motor expertise did not influence action understanding responses. Perhaps all performing musicians in our demography share a strong degree of understanding of the actions involved in piano-playing. As hypothesized, we observed that visual information enhanced observers' action understanding and liking ratings. However, contrary to our hypothesis, visual information had only a slight effect on observers' felt arousal responses. We observed a significant negative relationship between action understanding and felt affect, arousal, and valence responses.

Observers' music training appeared not to shape their felt affect responses to the musical performance stimuli. This suggests that the affect experienced in response to the performance of unfamiliar classical music compositions might occur independently of music training (Bigand et al., 2005). Indeed, a detailed analysis of such affective responses to four 20th century pieces showed that between-individual variations in these responses were far greater than inter-expertise group differences (Dean, Bailes, & Dunsmuir, 2014). However, some suggestions have been made that observers' musical experience might influence their neuro-affective responses induced by the performance of classical music compositions (Mikutta, Maissen, Altorfer, Strik, & König, 2014; Park et al., 2014). Nevertheless, yet other research suggests that individual differences in observers'

personality and musical preferences (Ladinig & Schellenberg, 2012), rather than their musical education (Grewe, Kopiez, & Altenmüller, 2009), might be linked more tightly with their affective responses to music. The complexity of the music presented here might have exerted an influence on observers' felt affect, particularly arousal, in a similar fashion regardless of their familiarity or previous experience with similar music (Marin, Lampatz, Wandl, & Leder, 2016).

Our results suggest that when the musical stimuli are complex and unfamiliar, observers' affective responses appear to be influenced predominantly by factors other than music training. This is in accord with the FEELA (force-effort-energy-loudness-affect) hypothesis (Dean & Bailes, 2010b; Olsen & Dean, 2016), which suggests that most listeners perceive a chain of influence from physical inputs to a musical sound through to loudness and affect. This is suggested to be largely independent of musical expertise or culture, not to require seeing a performer, and is likely one amongst many components of the action understanding we monitor in this work. A follow-up study will, therefore, investigate how characteristics of the stimuli, such as pitch, intensity, temporal, and timbral attributes of sound (Cespedes-Guevara & Eerola, 2018; Juslin & Laukka, 2003), or movement quantity and velocity (Davidson, 1994; Nusseck & Wanderley, 2009; Thompson & Luck, 2012) might explain observers' felt affect responses. Further research is also needed to investigate the degree to which an emotional contagion or empathy mechanism, perhaps underpinned by activation of a common neural network between performer and observer regardless of specific musical or motor training (Juslin, 2013; Juslin & Västfjäll, 2008), might be responsible for observers' felt affect responses to musical stimuli such as observed here. As no differences in arousal and valence responses between expertise groups were observed, it was deemed unnecessary to model liking and familiarity with arousal and valence further for the purposes of this study (Schubert, 2007).

Musical expertise did shape our observers' ability to cognitively take the perspective of the performer by understanding their (expressive) actions (in agreement with Egermann & McAdams, 2013). Shared embodied representations appeared to shape communication and understanding of expressive goal-directed actions between performer and observer (Corradini & Antonietti, 2013; Gallese, 2003). However, observers' action understanding responses appeared to be less-dependent on highly specific motor expertise (e.g., piano vs. non-piano) and shared perception-action networks with the performer (Calvo-Merino et al., 2005; Haslinger et al., 2005), and more related to a broader basis of shared

embodied experiences. Although speculative, the similarity of reported action understanding responses between the two musician groups might indicate some similar activation of a shared motor-related neural network, which potentially involves the temporal, fronto-parietal mirror neuron and limbic systems (Milston et al., 2013; Overy & Molnar-Szakacs, 2009). However, further research is needed to reveal the specific mechanisms underpinning observers' action understanding ratings. Our results suggest that cognitive perspective-taking with the performer and understanding their expressive action would appear to relate more to observers' expertise-moderated mental representations and models, which direct attention and influence cognitive decision-making processes about broader cues in the environment (Bellenkes et al., 1997; Broughton & Davidson, 2014; Broughton & Stevens, 2012; Doane et al., 2004; Schriver et al., 2008), than to the performer's specific instrument-playing actions.

The presence of visual information influenced affective and cognitive judgments in different ways. Although visual information has been shown to impact subjective and physiological measures of felt affect (Chapados & Levitin, 2008; Vines et al., 2006), contrary to our expectation the modality of presentation did not affect observers' valence responses, and visual information had only a small effect on observers' change in arousal responses. This suggests that arousal responses might be more influenced by bottom-up perceptual processing than valence responses (Kensinger & Corkin, 2004). Neuroscientific evidence supports the idea that arousal and valence might operate via distinct neural networks (e.g., Anders, Lotze, Erb, Grodd, & Birbaumer, 2004; Colibazzi et al., 2010; Gianotti et al., 2008). However, by and large in this study, observers' felt affect responses were invoked similarly through audio-only and audio-visual modes of presentation (Vuoskoski et al., 2016). It is possible that the response task might account for the unexpected results. The load placed on observers to monitor and report their felt affect on two dimensions simultaneously is arguably greater than having to report on one dimension (e.g., tension, Vines et al., 2006), or have physiological measures taken (Chapados & Levitin, 2008). It is possible that observers were operating at or close to cognitive capacity when making responses concurrently on two dimensions to auditory musical stimuli; the addition of visual stimuli might therefore have had no, or little, effect on their responses. In comparison, the presence of visual information coupled with the auditory enhanced observers' one-dimensional continuous action understanding and retrospective liking ratings

in this study. Future research is needed to examine cognitive load and multi-dimensional responses to multimodal music performance. However, the results of this study suggest that irrespective of musical expertise, when the performance is in a projected, public style, being able to see and hear the performing musician enhances observers' ability to cognitively take the musician's perspective and understand their expressive action, and their preference for the musical performance (Broughton & Stevens, 2009; Davidson, 1993; Schutz, 2008).

We predicted that musically trained observers would report higher liking and familiarity ratings than non-musician observers and found this effect most strong for liking ratings. It is plausible that musically trained observers would have prior exposure to and experience with similar musical stimuli, or the task of judging musical performance (Broughton & Stevens, 2009), which might have increased their familiarity with and preference for the stimuli (Zajonc, 2001). This idea is supported by our evidence that the two musician groups reportedly preferred more complex genres of music (e.g., *reflective and complex* and *intense and rebellious* dimensions) than the nonmusicians, who preferred genres of music in the *energetic and rhythmic* dimension, which is arguably less complex. Musician pianists reported their highest preference for the *reflective and complex* dimension of music, perhaps indicating a higher degree of familiarity with the most complex genres of music, given that the piano (unlike many instruments) is almost always a polyphonic instrument. It is also plausible that the pianists were more familiar with the genre of early 20th century classical solo piano music, if not the actual pieces performed, given our efforts to select music that would be unfamiliar, and that this increased their familiarity ratings in relation to the other musician group.

In this study, increased action understanding led to decreased arousal and valence responses, indicating that cognitive and affective systems are co-active in responding to musical performance. Potentially, increased understanding reflects the perceiver's experience and enhanced processing fluency (Reber, Schwarz, & Winkielman, 2004; Winkielman, Schwarz, Fazendeiro, & Reber, 2003), which has been proposed to lead to more positive affective responses. However, perhaps because the musical stimuli were complex and unfamiliar, increased action understanding reduced the subjective complexity and arousal potential of the stimuli leading to a reduction in felt-arousal responses (Berlyne, 1971, 1974). This idea is similar to Vuoskoski et al.'s (2016) suggestion that greater predictability of

events unfolding in the music might decrease arousal (in their study the audio-visual modality of presentation reduced skin conductance levels in comparison to audio-only), consistent with classic ideas of Meyer (and cf. Huron, 2006) on expectation. However, caution is also advised when interpreting the results of the change in arousal and valence models as the variance explained was modest. Furthermore, it is beyond the extent of current techniques to analyze how the audio-visual information fixed effect observed in the change in arousal model might be related to action understanding, as revealed as a significant effect in the initial modeling. The significant effect of time in the change in arousal model might reflect observers' experiencing some cognitive fatigue from sustained attention to the tasks throughout the session, which enhanced arousal (Head et al., 2016), or might reflect their increasing familiarity with the styles of the music.

The results of the present study should be considered in light of certain limitations. Most obviously, like most behavioral experiments ours creates demands (for indicating affective responses and action understanding) that may not always be part of participants' normal responses when listening to or viewing musical performance: this may depend in turn on their background and expertise. Thus, observers were assigned to the three different expertise groups according to their self-reported musical background. It is possible that there was some overlap in piano-playing expertise between the two musically trained groups. Future research should include objective assessment of musical and instrumental expertise through practical or standardized assessment tools. In addition, the measure of cognitive perspective taking in the form of action understanding warrants further consideration. In reporting action understanding, observers' scope of attention could have been highly varied, ranging from sound producing actions through to holistic bodily gesture, or even beyond to the task of performing for an audience. (A performer's actions commonly combine the necessary gestures involved in playing their instrument, with others that may relate either to their expressive intent, or to their own changing internal affect, from performance anxiety to affective responses to the present and imminent music.) The definition of action understanding could be refined to ensure that observers are interpreting the measure in the same manner, or the stimuli could be manipulated to direct attention to certain features and occlude others. The random effects by piece we observed in the action understanding, liking, familiarity, and change in arousal models suggests that characteristics of the music and observers' interactions with it

appear to have played a role in their affective and cognitive responses. The music in this study was unfamiliar, complex classical music and performed by a single, male pianist. Future research is needed to understand how observers' affective and cognitive responses might be influenced by varying attributes of the musical performance, such as musical complexity, using different performers, and individual differences in observers' experiences, personality, and preferences. A future study that compares affective and cognitive responses to music of varying combinations of human/machine creation/performance might help to tease apart the contribution of the performer versus the piece to observers' responses. It is already well known that acousmatic music, which is composed for presentation through loudspeakers alone and does not require a performer, can be affective (e.g., Bailes & Dean 2012). For the predictable future, however, both machine-generated and acousmatic music will still bear strong imprints of human creative processes, and these often also reflect performative processes. Future research should also aim to include objective measures of affective and cognitive responses through use of physiological, neuroimaging, and motion-capture tools to complement subjective self-report methods as used here.

IDEAS OF OBSERVER-PERFORMER EMPATHY IN MUSICAL INTERACTIONS

The present study drew on (and began to dissect) the idea that there exists a social connection between observer and performer in musical interactions, which facilitates empathic processes (Wöllner, 2017). Indeed, empathy is a key facet of our social worlds and our interactions with others (Davis, 1983, 1994) and musical performance is a context of social interaction. Empathy is a facet of a broader theory of embodied social cognition, which posits that others' intentions are manifest in expressive bodily activity and understood through shared motor, perceptual, and emotional experiences (Hostetter, Alibali, & Niedenthal, 2012). Felt affect, action understanding, and observer expertise are all likely involved to a degree in empathy responses. However, a thorough understanding of empathy in the context of music performance requires much further investigation to clarify the mechanisms and processes involved and how they interact.

Key mechanisms proposed to be involved in empathy include shared representations between actor and observer, underpinned by some common perception-action coding, and simulation or resonance mechanisms moderated by regulatory processing and self-other awareness (Decety & Jackson, 2004). Perception and

action are linked in that perception of another individual's behavior is proposed to automatically activate the observer's motor representation for that behavior (Preston & de Waal, 2002; Prinz, 1997).

This suggests that a critical question concerning affective aspects of observer empathy while experiencing music is whether the observers experience the same affective responses as the performer. And this might also be more specifically put: which aspects of the performer affect are shared? If there is sharing, it could be predominantly of affect related to the musical piece in question, but it might also to varying degrees involve sharing of anxiety, concentration or specific distractions, anticipation, feelings of success, etc., that a performer may experience. Such data do not seem to be available as yet. So sharing of affect may be a complex issue (and indeed it could not be measured here): it is proposed also to involve some degree of shared-representation and resonance or simulation mechanism, which includes coordinated autonomic and somatic responses. In a limited sense, affect sharing might be considered as emotional contagion—an automatic mimicry and synchronization of bodily movement and posture, vocal, and facial expressions with another to arrive at a similar emotional state (Hatfield, Rapson, & Le, 2011).

Such interpersonal emotional contagion is sometimes considered an unconscious process and a precursor to empathy. However, Egermann and McAdams (2013), in a large scale web study, operationalize emotional contagion differently: as the degree of parallel between perceptions of expressed affect (from one group of participants listening to five pieces amongst a wide range) and felt affect (from another group listening to five pieces from the same range). (A slight majority of participants were nonmusicians.) These two responses were quite similar to each other. All participants were then asked to evaluate the degree to which they could “empathize with the musicians you just heard” (p. 144) with no description or definition of empathy apparently provided. The empathy ratings were positive predictors of the degree of similarity between the expressed and felt affect values of a piece. Thus, the authors conclude that empathy mediates this form of emotional contagion, which rather than contagion between people, is arguably between aspects of a piece. The consideration of empathy between observer and performer thus has at least as many layers of potentially conflicting complexity as does social empathy in any other context.

Singer and Lamm (2009) nevertheless suggest that empathy is preceded by mimicry or emotional contagion, and followed by feelings of sympathy and

compassion, which might then lead to prosocial behavior. Surveying research on the neuroscience of empathy, Zaki and Ochsner (2012) offer three broad classifications of empathy processes: experience sharing, mentalizing or taking the perspective of another, and prosocial concern or motivation to improve the experiences of the other. As with the more limited parameters measured in our study (felt affect and action understanding; any ensuing prosocial concern or action tendencies were not the focus) these comprise both affective and cognitive components. There are probably both distinct and overlapping neural pathways for affective and cognitive components of empathy (i.e., dorsal mid-cingulate cortex for cognitive-evaluative and anterior insula for both cognitive and affective-perceptual forms of empathy; Fan et al., 2011) and the two forms are often coactive in natural (complex) social circumstances (Zaki & Ochsner, 2012).

The involvement of both affective and cognitive components in empathy suggests equally the possible contribution of both bottom-up and top-down processes. Thus, humans' use of visual and auditory signals to affectively empathize with others (Fan et al., 2011; Warren et al., 2006) may involve both bottom-up emotional contagion and top-down appraisal processes (Preston & de Waal, 2002; Singer & Lamm, 2009). It should also be noted that although observers might be able to perspective take and understand another's expressive action, that does not necessarily mean that they will empathize with the other person. There might well be other processes involved, such as regulation (Decety & Jackson, 2004).

Many questions thus remain unresolved in relation to social empathy at large, and its specific relevance to musical appreciation in particular.

CONCLUSIONS

Musical performance represents a context of social interaction where thoughts and feelings can be shared between performers and observers. The results of the present study support the notion that empathy and embodied social cognitive theory might apply to musical performance, even if with considerable complexity. Shared embodied experiences between observer and performer appear to be important for observers to connect with and understand the performer through their expressive action. Our findings indicate that whereas musical (but not specialized motor) expertise, and modality of presentation appear to influence observers' cognitive responses, affective responses appear to be robust against variations in modality of presentation or observers' musical or specific motor expertise. The

framework presented here assists conceptualizing how observers with different backgrounds connect with performers, and how affective and cognitive responses are related. The results of this study suggest that when observers are faced with musical performance that is cognitively challenging, their experience with and mental representations of similar stimulus and environments appears to influence the degree to which they can connect with the performer, understand what they're doing and their preferences for the music. New strategies to motivate and develop audiences for less familiar and more cognitively challenging musical performance might usefully be based on developing observers' understanding of the musician in the act of performing. Such strategies might assist in developing new audiences for more challenging musical performance practices, and work as a complement to traditional marketing approaches (Barlow & Shibli, 2007; Kolb, 2013).

Author Note

This work was supported by The University of Queensland Early Career Researcher Award [grant number 2014003045] granted to the first author.

Ethical approval for this project was given by The University of Queensland Behavioural & Social Sciences Ethical Review Committee [approval number 2015000462].

Data can be obtained by emailing the corresponding author.

The Author(s) declare(s) that there is no conflict of interest.

Correspondence concerning this article should be addressed to Mary C. Broughton, Psychology of Music Program Convener, School of Music, The University of Queensland, Level 4, Zelman Cowen Building, St Lucia, Brisbane, Queensland 4072, Australia. E-mail: m.broughton@uq.edu.au

References

- ANDERS, S., LOTZE, M., ERB, M., GRODD, W., & BIRBAUMER, N. (2004). Brain activity underlying emotional valence and arousal: A response-related fMRI study. *Human Brain Mapping, 23*, 200–209. DOI: 10.1002/hbm.20048
- BAILES, F., & DEAN, R. T. (2012). Comparative time series analysis of perceptual responses to electroacoustic music. *Music Perception, 29*, 359–375. DOI: 10.1525/mp.2012.29.4.359
- BALCH, W. R., & LEWIS, B. S. (1996). Music-dependent memory: The roles of tempo changes and mood mediation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1354–1363. DOI: 10.1037/0278-7393.22.6.1354
- BARLOW, M., & SHIBLI, S. (2007). Audience development in the arts: A case study of chamber music. *Managing Leisure, 12*, 102–119. DOI: 10.1080/13606710701339272
- BARON-COHEN, S., WHEELWRIGHT, S., SKINNER, R., MARTIN, J., & CLUBLEY, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17. DOI: 10.1023/A:1005653411471
- BELLENKES, A. H., WICKENS, C. D., & KRAMER, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine, 68*, 569–579. PMID: 9215461
- BERLYNE, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
- BERLYNE, D. E. (Ed.). (1974). *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Oxford, England: Hemisphere.
- BIGAND, E., VIEILLARD, S., MADURELL, F., MAROZEAU, J., & DACQUET, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion, 19*, 1113–1139.
- BROUGHTON, M. C., & DAVIDSON, J. W. (2014). Action and familiarity effects on self and other expert musicians' Laban effort-shape analyses of expressive bodily behaviors in instrumental music performance: A case study approach. *Frontiers in Psychology, 5*, 1201. DOI: 10.3389/fpsyg.2014.01201
- BROUGHTON, M. C., & STEVENS, C. J. (2009). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music, 37*, 137–153. DOI: 10.1177/0305735608094511
- BROUGHTON, M. C., & STEVENS, C. J. (2012). Analyzing expressive qualities in movement and stillness: Effort-shape analyses of solo marimbists' bodily expression. *Music Perception, 29*, 339–357. DOI: 10.1525/mp.2012.29.4.339
- CALVO-MERINO, B., GLASER, D. E., GRÈZES, J., PASSINGHAM, R. E., & HAGGARD, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249. DOI: 10.1093/cercor/bhi007
- CARR, L., IACOBONI, M., DUBEAU, M. C., MAZZIOTTA, J. C., & LENZI, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of National Academy of Science USA, 100*, 5497–5502. DOI: 10.1073/pnas.0935845100

- CESPEDES-GUEVARA, J., & EEROLA, T. (2018). Music communicates affects, not basic emotions—a constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology, 9*, 215. DOI: 10.3389/fpsyg.2018.00215
- CHAPADOS, C., & LEVITIN, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition, 108*, 639–651. DOI: 10.1016/j.cognition.2008.05.008
- CHMIEL, A., & SCHUBERT, E., (2019). Psycho-historical contextualization for music and visual works: A literature review and comparison between artistic mediums. *Frontiers in Psychology, 10*, 182. DOI: 10.3389/fpsyg.2019.00182
- COLIBAZZI, T., POSNER, J., WANG, Z., GORMAN, D., GERBER, A., YU, S., ET AL. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion, 10*, 377–389. DOI: 10.1037/a0018484
- CORRADINI, A., & ANTONIETTI, A. (2013). Mirror neurons and their function in cognitively understood empathy. *Consciousness and Cognition, 22*, 1152–1161. DOI: 10.1016/j.concog.2013.03.003
- DAVIDSON, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21*, 103–113. DOI: 10.1177/030573569302100201
- DAVIDSON, J. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies, 26*, 279–301.
- DAVIDSON, J. W., & EDGAR, R. (2003). Gender and race bias in the judgement of Western art music performance. *Music Education Research, 5*, 169–181. DOI: 10.1080/1461380032000085540
- DAVIS, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113–126. DOI: 10.1037/0022-3514.44.1.113
- DAVIS, M. H. (1994). *Empathy: A social psychological approach*. Madison, WI: Brown & Benchmark.
- DEAN R. T., & BAILES, F. (2010a) Time series analysis as a method to examine acoustical influences on real-time perception of music. *Empirical Musicology Review, 5*, 152–175.
- DEAN, R. T., & BAILES, F.A. (2010b). The control of acoustic intensity during jazz and free improvisation performance. *Critical Studies in Improvisation/Études critiques en improvisation, 6*(2), 1–22.
- DEAN, R. T., BAILES, F., & DUNSMUIR, W. T. M. (2014). Shared and distinct mechanisms of individual and expertise-group perception of expressed arousal in four works. *Journal of Mathematics and Music: Mathematical and Computational Approaches to Music Theory, Analysis, Composition and Performance, 8*, 207–223. DOI: 10.1080/17459737.2014.928753
- DEAN, R. T., BAILES, F., & SCHUBERT, E. (2011). Acoustic intensity causes perceived changes in arousal levels in music: An experimental investigation. *PLoS ONE, 6*, e18591. DOI: 10.1371/journal.pone.0018591
- DEAN R. T., & DUNSMUIR W. T. (2016) Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods, 48*, 783–802. DOI: 10.3758/s13428-015-0611-2
- DEARN, L. K., & PRICE, S. M. (2016). Sharing music: Social and communal aspects of concert-going. *Networking Knowledge: Journal of the MeCCSA Postgraduate Network, 9*(2), 1–20. DOI: 10.31165/nk.2016.92.428
- DECETY, J., & JACKSON, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*, 71–100. DOI: 10.1177/1534582304267187
- DECETY, J., & LAMM, C. (2009). The biological basis of empathy and intersubjectivity. In J. T. Cacioppo & G. G. Berntson (Eds.), *Handbook of neuroscience for the behavioral sciences* (pp. 940–957). New York: John Wiley and Sons.
- DOANE, S. M., SOHN, Y. W., & JODŁOWSKI, M. T. (2004). Pilot ability to anticipate the consequences of flight actions as a function of expertise. *Human Factors, 46*, 92–103. DOI: 10.1518/hfes.46.1.92.30386
- DOBSON, M. C. (2010). New audiences for classical music: The experiences of non-attenders at live orchestral concerts. *Journal of New Music Research, 39*, 111–124. DOI: 10.1080/09298215.2010.489643
- DOBSON, M. C., & PITTS, S. E. (2011). Classical cult or learning community? Exploring new audience members' social and musical responses to first-time concert attendance. *Ethnomusicology Forum, 20*, 353–383. DOI: 10.1080/17411912.2011.641717
- EEROLA, T., & VUOSKOSKI, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception, 30*, 307–340. DOI: 10.1525/mp.2012.30.3.307
- EGERMANN, H., & McADAMS, S. (2013). Empathy and emotional contagion as a link between recognized and felt emotions in music listening. *Music Perception, 31*, 139–156. DOI: 10.1525/mp.2013.31.2.139
- EGERMANN, H., PEARCE, M. T., WIGGINS, G. A., & McADAMS, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, and Behavioral Neuroscience, 13*, 533–553. DOI: 10.3758/s13415-013-0161-y
- ELLIOT, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion, 30*, 111–116. DOI: 10.1007/s11031-006-9028-7
- 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. DOI: 10.1037/0033-295X.100.3.363

- FAN, Y., DUNCAN, N. W., DE GRECK, M., & NORTHOFF, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews*, 35, 903–911. DOI: 10.1016/j.neubiorev.2010.10.009
- GALLESE, V. (2003). The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology*, 36, 171–180. DOI: 10.1159/000072786
- GAZZOLA, V., AZIZ-ZADEH, L., & KEYSERS, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current Biology*, 16, 1824–1829. DOI: 10.1016/j.cub.2006.07.072
- GIANOTTI, L. R., FABER, P. L., SCHULER, M., PASCUAL-MARQUI, R. D., KOCHI, K., & LEHMANN, D. (2008). First valence, then arousal: The temporal dynamics of brain electric activity evoked by emotional stimuli. *Brain Topography*, 20, 143–156. DOI: 10.1007/s10548-007-0041-2
- GREWE, O., KOPIEZ, R., & ALTENMÜLLER, E. (2009). The chill parameter: Goose bumps and shivers as promising measures in emotion research. *Music Perception*, 27, 61–74. DOI: 10.1525/mp.2009.27.1.61
- GREWE, O., NAGEL, F., KOPIEZ, R., & ALTENMÜLLER, E. (2007). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion*, 7, 774–788. DOI: 10.1037/1528-3542.7.4.774
- HASLINGER, B., ERHARD, P., ALTENMÜLLER, E., SCHROEDER, U., BOECKER, H., & CEBALLOS-BAUMANN, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, 17, 282–293. DOI: 10.1162/0898929053124893
- HATFIELD, E., RAPSON, R. L., & LE, Y. C. L. (2011). Emotional contagion and empathy. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 19–30). Cambridge, MA: MIT Press.
- HEAD, J. R., TENAN, M. S., TWEDELL, A. J., PRICE, T. F., LAFIANDRA, M. E., & HELTON, W. S. (2016). Cognitive fatigue influences time-on-task during bodyweight resistance training exercise. *Frontiers in Physiology*, 7, 373. DOI: 10.3389/fphys.2016.00373
- HILLS, A. (2003). *Foolproof guide to statistics using SPSS*: Sydney, Australia: Pearson Education.
- HOSTETTER, A. B., ALIBALI, M. W., & NIEDENTHAL, P. M. (2012). *Embodied social thought: Linking social concepts, emotion, and gesture*. London, England: SAGE Publications.
- HUSAIN, G., THOMPSON, W. F., & SCHELLENBERG, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, 20, 151–171. DOI: 10.1525/mp.2002.20.2.151
- HURON, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- ICKES, W. J. (Ed.). (1997). *Empathic accuracy*. New York: Guilford Press.
- JUCHNIEWICZ, J. (2008). The influence of physical movement on the perception of musical performance. *Psychology of Music*, 36, 417–427. DOI: 10.1177/0305735607086046
- JUSLIN, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10, 235–266. DOI: 10.1016/j.plrev.2013.05.008
- JUSLIN, P. N., & LAUKKA, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129, 770–814. DOI: 10.1037/0033-2909.129.5.770
- JUSLIN, P. N., & VÄSTFJÄLL, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31, 559–575. DOI: 10.1017/S0140525X08005293
- KEMP, E., & WHITE, M. G. (2013). Embracing jazz: Exploring audience participation in jazz music in its birthplace. *International Journal of Arts Management*, 16, 35–81.
- KENSINGER, E. A., & CORKIN, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, 101, 3310–3315. DOI: 10.1073/pnas.0306408101
- KOLB, B. M. (2013). *Marketing for cultural organizations: New strategies for attracting and engaging audiences*. New York: Routledge.
- KRAMPE, R. T., & ERICSSON, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, 125, 331–359. DOI: 10.1037/0096-3445.125.4.331
- LADINIG, O., & SCHELLENBERG, E. G. (2012). Liking unfamiliar music: Effects of felt emotion and individual differences. *Psychology of Aesthetics, Creativity, and the Arts*, 6, 146–154. DOI: 10.1037/a0024671
- LEVINSON, J. (2011). *Music, art, and metaphysics. Essays in philosophical aesthetics*. Oxford, UK: Oxford University Press.
- LOMBARDO, M. V., BARNES, J. L., WHEELWRIGHT, S. J., & BARON-COHEN, S. (2007). Self-referential cognition and empathy in autism. *PLoS ONE*, 2, 883. DOI: 10.1371/journal.pone.0000883
- MARIN, M. M., LAMPATZ, A., WANDL, M., & LEDER, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in Human Neuroscience*, 10, 536. DOI: 10.3389/fnhum.2016.00536
- MCNEILL, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- MIKUTTA, C. A., MAISSEN, G., ALTORFER, A., STRIK, W., & KÖNIG, T. (2014). Professional musicians listen differently to music. *Neuroscience*, 268, 102–111. DOI: 10.1016/j.neuroscience.2014.03.007

- MILSTON, S. I., VANMAN, E. J., & CUNNINGTON, R. (2013). Cognitive empathy and motor activity during observed actions. *Neuropsychologia*, *51*, 1103–1108. DOI: 10.1016/j.neuropsychologia.2013.02.020
- MIU, A. C., & BALTES, F. R. (2012). Empathy manipulation impacts music-induced emotions: A psychophysiological study on opera. *PLoS ONE*, *7*, e30618. DOI: 10.1371/journal.pone.0030618
- MIU, A. C., & VUOSKOSKI, J. K. (2017). The social side of music listening: Empathy and contagion in music-induced emotions. In E. King & C. Waddington (Eds.), *Music and empathy* (pp. 124–138). London, UK: Routledge.
- NAGEL, F., KOPIEZ, R., GREWE, O., & ALTENMÜLLER, E. (2007). EMuJoy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods*, *39*, 283–290. DOI: 10.3758/BF03193159
- NUSSECK, M., & WANDERLEY, M. M. (2009). Music and motion – How music-related ancillary body movements contribute to the experience of music. *Music Perception*, *26*, 335–353. DOI: 10.1525/mp.2009.26.4.335
- OLSEN, K. N., & DEAN, R. T. (2016). Does perceived exertion influence perceived affect in response to music? Investigating the “FEELA” hypothesis. *Psychomusicology: Music, Mind, and Brain*, *26*(3), 257–269.
- OVERY, K., & MOLNAR-SZAKACS, I. (2009). Being together in time: Musical experience and the Mirror Neuron System. *Music Perception*, *26*, 489–504. DOI: 10.1525/mp.2009.26.5.489
- PARK, M., GUTYRCHIK, E., BAO, Y., ZAYTSEVA, Y., CARL, P., WELKER, L., ET AL. (2014). Differences between musicians and nonmusicians in neuro-affective processing of sadness and fear expressed in music. *Neuroscience Letters*, *566*, 120–124. DOI: 10.1016/j.neulet.2014.02.041
- PITTS, S. E. (2005). What makes an audience? Investigating the roles and experiences of listeners at a chamber music festival. *Music and Letters*, *86*, 257–269. DOI: 10.1093/ml/gci035
- PRESTON, S. D., & DE WAAL, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioural and Brain Sciences*, *25*, 1–20. DOI: 10.1017/S0140525X02000018
- PRINZ, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154. DOI: 10.1080/713752551
- RADBOURNE, J., JOHANSON, K., GLOW, H., & WHITE, T. (2009). The audience experience: Measuring quality in the performing arts. *International Journal of Arts Management*, *11*, 16–29. DOI: 10.2307/41064995
- REBER, R., SCHWARZ, N., & WINKIELMAN, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver’s processing experience? *Personality and Social Psychology Review*, *8*, 364–382. DOI: 10.1207/s15327957pspr0804_3
- RENTFROW, P. J., & GOSLING, S. D. (2003). The do re mi’s of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*, 1236–1256. DOI: 10.1037/0022-3514.84.6.1236
- RUSSELL, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178. DOI: 10.1037/h0077714
- SHELLEKENS, E., & GOLDIE, P. (2011). *The aesthetic mind: Philosophy and psychology*. Oxford, UK: Oxford University Press.
- SCHERER, K. R., & COUTINHO, E. (2013). How music creates emotion: A multifactorial process approach. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The emotional power of music: Multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 121–145). Oxford, UK: Oxford University Press.
- SCHRIVER, A. T., MORROW, D. G., WICKENS, C. D., & TALLEUR, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. DOI: 10.1518/001872008X374974
- SCHUBERT, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, *51*, 154–165. DOI: 10.1080/00049539908255353
- SCHUBERT, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, *21*, 561–585. DOI: 10.1525/mp.2004.21.4.561
- SCHUBERT, E. (2006). Analysis of emotional dimensions in music using time series techniques. *Context*, *31*, 65–80.
- SCHUBERT, E. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, *35*, 499–515. DOI: 10.1177/0305735607072657
- SCHUTZ, M. (2008). Seeing music? What musicians need to know about vision. *Empirical Musicology Review*, *3*, 83–108. DOI: 10.18061/1811/34098
- SCHUTZ, M., & LIPSCOMB, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*, 888–897. DOI: 10.1068/p5635
- SHAFIR, T., TAYLOR, S. F., ATKINSON, A. P., LANGENECKER, S. A., & ZUBIETA, J. K. (2013). Emotion regulation through execution, observation, and imagery of emotional movements. *Brain and Cognition*, *82*, 219–227. DOI: 10.1016/j.bandc.2013.03.001
- SHAFIR, T., TSACHOR, R. P., & WELCH, K. B. (2016). Emotion regulation through movement: Unique sets of movement characteristics are associated with and enhance basic emotions. *Frontiers in Psychology*, *6*, 2030. DOI: 10.3389/fpsyg.2015.02030
- SINGER, T., & LAMM, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, *1156*, 81–96. DOI: 10.1111/j.1749-6632.2009.04418.x

- SIZOO, B. B., HORWITZ, E., TEUNISSE, J., KAN, C., VISSERS, C., FORCEVILLE, E., ET AL. (2015). Predictive validity of self-report questionnaires in the assessment of autism spectrum disorders in adults. *Autism, 19*, 842–849. DOI: 10.1177/1362361315589869
- STEVENS, C., VINCS, K., & SCHUBERT, E. (2009). Measuring audience response on-line: An evaluation of the portable Audience Response Facility (pARF). In C. Stevens, E. Schubert, B. Kruithof, K. Buckley, & S. Fazio (Eds.), *Proceedings of the 2nd International Conference on Music Communication Science (ICoMCS2)*. Sydney, Australia: HCSNet, Western Sydney University.
- STEVENS, C. J., DEAN, R. T., VINCS, K., & SCHUBERT, E. (2014). In the heat of the moment: Audience real-time response to music and dance performance. In K. Burland & S. Pitts (Eds.), *Coughing and clapping: Investigating audience experience* (pp. 69–87). Retrieved from <http://UWSAU.ebib.com.au/patron/FullRecord.aspx?p=1815567>
- TAJTÁKOVÁ, M., & ARIAS-ARANDA, D. (2008). Targeting university students in audience development strategies for opera and ballet. *The Service Industries Journal, 28*, 179–191. DOI: 10.1080/02642060701842191
- THOMPSON, M. R., & LUCK, G. (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae, 16*, 19–40. DOI: 10.1177/1029864911423457
- VINES, B. W., KRUMHANSL, C. L., WANDERLEY, M. M., & LEVITIN, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition, 101*, 80–113. DOI: 10.1016/j.cognition.2005.09.003
- VUOSKOSKI, J. K., GATTI, E., SPENCE, C., & CLARKE, E. F. (2016). Do visual cues intensify the emotional responses evoked by musical performance? A psychophysiological investigation. *Psychomusicology: Music, Mind, and Brain, 26*, 179–188. DOI: 10.1037/pmu0000142
- WALMSLEY, B. (2011). Why people go to the theatre: A qualitative study of audience motivation. *Journal of Customer Behaviour, 10*, 335–351. DOI: 10.1362/147539211X13210329822545
- WARREN, J. E., SAUTER, D. A., EISNER, F., WILAND, J., DRESNER, M. A., WISE, R. J., ET AL. (2006). Positive emotions preferentially engage an auditory-motor “mirror” system. *Journal of Neuroscience, 26*, 13067–13075. DOI: 10.1523/jneurosci.3907-06.2006
- WINKIELMAN, P., SCHWARZ, N., FAZENDEIRO, T., & REBER, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- WÖLLNER, C. (2017). Audience responses in the light of perception-action theories of empathy. In E. King & C. Waddington (Eds.), *Music and empathy* (pp. 139–156). London, UK: Routledge.
- WÖLLNER, C., & CAÑAL-BRULAND, R. (2010). Keeping an eye on the violinist: Motor experts show superior timing consistency in a visual perception task. *Psychological Research, 74*, 579–585. DOI: 10.1007/s00426-010-0280-9
- YULE, G. U. (1926). Why do we sometimes get nonsense- correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society, 89*, 1–63.
- ZAJONC, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*, 224–228. DOI: 10.1111/1467-8721.00154
- ZAKI, J., & OCHSNER, K. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15*, 675–680. DOI: 10.1038/nn.3085
- ZHANG, J. D., & SCHUBERT, E. (2019). A single item measure for identifying musician and nonmusician categories based on measures of musical sophistication. *Music Perception, 36*, 457–467. DOI: 10.1525/mp.2019.36.5.457