Article

Psychophysiological Response to Acoustic Intensity Change in a Musical Chord

Kirk N. Olsen and Catherine J. Stevens

MARCS Institute and School of Social Sciences and Psychology, University of Western Sydney, Penrith, NSW, Australia

Abstract. This paper investigates psychological and psychophysiological components of arousal and emotional response to a violin chord stimulus comprised of continuous increases (up-ramp) or decreases (down-ramp) of intensity. A factorial experiment manipulated direction of intensity change (60–90 dB SPL up-ramp, 90–60 dB SPL down-ramp) and duration (1.8 s, 3.6 s) within-subjects ($N = 45$). Dependent variables were ratings of emotional arousal, valence, and loudness change, and a fine-grained analysis of event-related skin conductance response (SCR). As hypothesized, relative to down-ramps, musical up-ramps elicited significantly higher ratings of emotional arousal and loudness change, with marginally longer SCR rise times. However, SCR magnitude was greater in response to musical down-ramps. The implications of acoustic intensity change for music-induced emotion and auditory warning perception are discussed.

Keywords: arousal, auditory looming, emotion, loudness change, music, skin conductance

Music is intimately and profoundly associated with emotion (Juslin & Sloboda, 2001). Emotions induced by music are related to broader motivational determinants of emotion, such as arousal and valence (Khalfa, Peretz, Blondin, & Manon, 2002; Russell, 1980). Music-induced emotions are associated with physiological indices of autonomic arousal (Lang, Bradley, & Cuthbert, 1998) and vary with acoustic parameters such as musical mode (e.g., major or minor key), tempo (e.g., Dalla Bella, Peretz, Rousseau, & Gosselin, 2001; Gagnon & Peretz, 2003; Heinlein, 1928), texture (e.g., Gabrielsson & Lindström, 2001; Sloboda, 1991), and intensity (Balkwill & Thompson, 1999; Balkwill, Thompson, & Matsunaga, 2004; Krumhansl, 1997; Schubert, 2004). Music is inherently dynamic, changing continuously as it unfolds through time. However, systematic empirical investigations of the temporal *dynamics* of acoustic variables such as continuous intensity change in music are scarce, especially concerning psychological, physiological, and perceptual indices of emotional response. The present study investigates self-report and psychophysiological measures of arousal elicited by a dynamic acoustic dimension: continuous intensity change embedded within a musical chord.

Human Response to Acoustic Intensity Change

In psychophysical terms, the percept of loudness is closely related to a sound's physical intensity. In auditory motion perception, a continuous increase of acoustic intensity (termed hereafter an "up-ramp") is a key indicator of a "looming" or approaching stimulus in the environment, whereas a continuous decrease of intensity (termed hereafter a "down-ramp") can indicate a receding sound source. When these dynamic stimuli are identical in terms of spectral frequency, duration, range, and region of intensity change, up-ramps are perceptually louder (Ries, Schlauch, & DiGiovanni, 2008; Stecker & Hafter, 2000; Susini, McAdams, & Smith, 2007), subjectively longer in duration (DiGiovanni & Schlauch, 2007; Grassi & Darwin, 2006; Ries et al., 2008; Schlauch, Ries, & DiGiovanni, 2001), and are judged to change more in loudness (Bach, Neuhoff, Perrig, & Seifritz, 2009; Neuhoff, 1998, 2001; Olsen, Stevens, & Tardieu, 2010; Seifritz et al., 2002).

This difference in perception – particularly in relation to the overestimation of loudness change in response to

DOI: 10.1027/0269-8803/a000082

pure-tone up-ramps – has led to the hypothesis of an adaptive perceptual bias to looming auditory motion (Neuhoff, 1998, 2001). That is, there may be an adaptive mechanism in response to a looming and potentially threatening stimulus. Such a mechanism would allow an organism extra time for survival behavior such as avoidance or retreat (Neuhoff, 2004). Visual analogs of looming exist; for example, human infants show fear responses to an expanding visual display that represents impending collision (Nanez, 1988), while human adults underestimate the time-to-contact of a looming visual stimulus, expecting contact significantly earlier than actual contact (Schiff & Oldak, 1990).

A "looming-specific" neural network activated in response to up-ramp stimuli has been hypothesized to direct attention to the location and direction of movement of a sound source, aiding rapid decision making and action (Bach et al., 2008; Hall & Moore, 2003; Maier & Ghazanfar, 2007; Seifritz et al., 2002). A "looming-specific" response is also associated with a heightened magnitude of psychophysiological arousal indicated by the human skin conductance response (SCR). This response is specific to short (2 s) up-ramp stimuli comprising full-motion cues (simulating a three-dimensional approaching sound source) and intensity change alone (Bach et al., 2009). A specific neural network and characteristic psychophysiological arousal response to up-ramp stimuli are also associated with significantly higher ratings of emotional arousal and greater negative valence (i.e., unpleasantness).

Neuroimaging, physiological, and psychological data suggest a fundamental response to the warning properties of looming auditory motion in general, and a continuous increase of acoustic intensity in particular. If there is a fundamental response to increasing intensity, then we would expect comparable results in real-world listening domains where continuous acoustic intensity change is common – for example, in music. Such a response may, in part, inform our understanding of psychophysiological and emotional arousal responses to music.

Psychophysiological Response to Acoustic Intensity in Music

Intensity is a salient acoustic dimension that acts as a reliable cue to music's intended emotion(s) across cultures (e.g., Western, Japanese, and Hindustani music) (Balkwill et al., 2004) and is positively correlated with the arousal dimension of Russell's (1980) circumplex arousal-valence model of emotion (Schubert, 2004). Peaks of loudness directly related to acoustic intensity have been linked to musically induced self-reported "chills" (Nagel, Kopiez, Grewe, & Altenmüller, 2008) that are associated with increases of psychophysiological arousal (e.g., the SCR) (Grewe, Nagel, Kopiez, & Altenmüller, 2007; Guhn, Hamm, & Zentner, 2007).

However, studies investigating intensity in music have often treated it as a static variable; that is, set to a high or low level (e.g., Ilie & Thompson, 2006). Such an approach overlooks the inherent dynamic properties of music that change through time. For example, musicological analyses of scores from the classical and romantic periods of Western tonal music have shown that musical crescendos – which are associated with increases of intensity – are more frequent and cover a greater duration of dynamic change than diminuendos (decreases of intensity), which are shorter in duration and proportionally less frequent (Huron, 1990, 1991, but see Dean & Bailes, 2010). The prevalence of gradual and extended increases of intensity from notated crescendos in musical scores – what Huron calls a "ramp archetype" – functions to maintain listeners' attention throughout a piece of music (Huron, 1992). Sustained attention elicited from up-ramp stimuli in a musical context may be underpinned by the hypothesized "looming-specific" neural network (Hall & Moore, 2003) and coincide with significant increases of arousal and loudness change in reports investigating nonmusical up-ramps (e.g., Bach et al., 2009; Neuhoff, 1998). Recent evidence from perceptual experiments using musical stimuli supports this hypothesis: loudness change is overestimated in response to short duration (1.8 s and 3.6 s) up-ramps using a violin timbre with varied stimulus complexity (e.g., a single-note stimulus and diminished triad chord; Olsen et al., 2010).

The rate of intensity change (dB SPL change/time) of a dynamic stimulus may also be a contributing factor to the psychophysiological arousal response to music. Continuous time-series reveal that the more sudden a change in loudness (e.g., 1–2 s compared to 2–3 s), the faster the change in self-reported arousal (Schubert & Dunsmuir, 1999). A faster rate of intensity change in crescendos of the music of Brahms and Scriabin is significantly correlated with self-reported increases of arousal and shivers down the spine (Yasuda, 2009). It follows, therefore, that this heightened experience of subjective arousal in response to a fast rate of increasing intensity should be associated with an accompanying psychophysiological response specific to that direction and rate of intensity change. If higher ratings of subjective arousal are correlated with a faster rate of intensity change, then we would expect a longer psychophysiological response of greater magnitude to a musical stimulus rapidly increasing in intensity, relative to slower rates of intensity change.

We argue that a fundamental psychophysiological response is elicited by up-ramp stimuli in general and should be recovered when investigated in a specific musical context. In conjunction with perceptual overestimation of intensity change, we predict that explicit ratings of emotional arousal, negative valence, and increases of psychophysiological response (SCR) are greater in response to up-ramps, relative to down-ramps, when presented as a musical (violin) timbre over two stimulus durations (1.8 s, 3.6 s). These predictions will be tested by measuring: (1) judged loudness change, (2) ratings of emotional arousal and valence, and (3) event-related SCR.

Event-Related SCR

As a measure of autonomic arousal that is sensitive to music-induced emotions (Khalfa et al., 2002), event-related SCR is a phasic and transient change in skin conductance in response to one short stimulus event (usually < 10 s). This contrasts with the measurement of tonic levels of skin conductance over long stimulus durations (many minutes) (e.g., Krumhansl, 1997). However, psychophysiological studies of music measuring skin conductance commonly report only SCR magnitudes or amplitudes, which is limiting because important temporal aspects of the event-related SCR are omitted.

The present study includes a more fine-grained analysis of the SCR. In conjunction with SCR magnitude and amplitude, the temporal measures of response latency and rise time will be analyzed (Andreassi, 2000; Venables & Christie, 1980). SCR latency can be conceptualized as the response lag: the time-frame between stimulus onset and SCR onset. A shorter SCR latency and thus quicker response onset would reflect the behavioral priority of an up-ramp stimulus. SCR rise time refers to the duration of each SCR, measured from response onset to response peak. This is a temporal measure of the SCR amplitude, where a longer rise time is evidence of a longer SCR.

To calculate SCR amplitude, latency, and rise time, the criterion for an event-related SCR was set as a 0.05 μS increase in skin conductance level (SCL) relative to SCL at stimulus onset (baseline). The temporal window for valid response onset was 1–4 s measured from stimulus onset. In other words, if the response criterion of a 0.05 μS increase in SCL (relative to baseline SCL) occurred between 1 s and 4 s after stimulus onset, it was deemed a valid SCR. SCR amplitude, latency, and rise time were then calculated for valid SCRs. If the 0.05 μS threshold was not reached, no response was coded. In addition, SCR magnitude was calculated as the mean SCL between 4 and 5 s after stimulus onset, corrected for a SCL baseline period of 1 s before stimulus onset (Bach et al., 2008, 2009). SCR magnitude results in a response value for each trial, even if the 0.05 μS criterion was not met. In sum, analyses of SCR amplitude, latency, and rise time data omit no-response trials from the analysis, whereas SCR magnitude includes all trials.

Aim, Design, and Hypotheses

The overarching aim of the experiment was to investigate the psychophysiological response (i.e., the SCR) to up-ramp and down-ramp musical stimuli. To complement the psychophysiological data, explicit ratings of emotional arousal and valence were recorded. As a psychological indicator of acoustic intensity change, judged loudness change was measured. The experiment was realized as a 2 × 2 within-subjects factorial design. All up-ramp and down-ramp musical stimuli comprised a sampled violin timbre presented as a four-note diminished chord (i.e., consisting of three simultaneous minor third (“sad”) intervals). Independent variables were the direction of intensity change (up-ramp, down-ramp) and stimulus duration (1.8 s, 3.6 s).

The psychological dependent variables were explicit ratings of arousal, valence, and loudness change. The psychophysiological dependent measures were SCR magnitude, SCR amplitude, SCR latency, and SCR rise time. Figure 1 shows an example of an SCR to a 3.6 s down-ramp stimulus with SCR amplitude, latency, and rise time labeled.

Specifically, we hypothesized that relative to down-ramps:

> *Hypothesis 1*: Up-ramps are overestimated in loudness change;
>
> *Hypothesis 2*: Up-ramps elicit higher ratings of arousal and greater negative valence (unpleasantness);
>
> *Hypothesis 3*: Up-ramps elicit shorter SCR latency, longer SCR rise time, and greater SCR magnitude.

Method

Participants

The sample consisted of 45 adult participants recruited from the University of Western Sydney (34 females and 11 males; *M* = 20.95 years, *SD* = 4.93, range = 18–46 years). All reported normal hearing. Sixteen participants had received minimal individual musical training (*M* = 1.69 years, *SD* = .79, range = 1–3 years).

Stimuli

Stimuli comprised a linear intensity increase (up-ramp) or decrease (down-ramp) from 60 to 90 dB SPL and 90 to 60 dB SPL, respectively. A linear change was chosen for methodological consistency with previous research using up-ramps and down-ramps (Bach et al., 2009; Neuhoff, 1998, 2001; Olsen et al., 2010). The generation of violin stimuli began with a 1.8 s and 3.6 s steady-state recorded violin sample (sampling frequency of 44.1 kHz). Each stimulus consisted of a dissonant diminished four-note chord structure, C_4 (F_0 = 261.63 Hz), $\mathrm{E}_4^{\mathrm{b}}$ (F_0 = 311.13 Hz), $\mathrm{G}_4^{\mathrm{b}}$ (F_0 = 369.99 Hz), and A_4 (F_0 = 440.00 Hz). Up-ramps and down-ramps were constructed from the steady-state exemplars in a sound-attenuated booth using a custom computer program written in MAX-MSP (Version 4.6.3). A minimum (60 dB SPL) and maximum (90 dB SPL) intensity level was recorded in the MAX-MSP program from each steady-state sound. The program generated an up-ramp and a down-ramp for each condition by using the two recorded dB SPL levels as onset/offset anchors and creating a linear intensity change between them using each original steady-state sound. Ten millisecond fade-in and fade-out ramps were incorporated to remove onset/offset artifacts in the stimuli (such as audible “clicks”). Two directions of intensity change crossed with two stimulus durations resulted in four stimuli overall.

Each of the four violin stimuli included variable durations of silence (range = 10–12 s) presented at the beginning and 1 s of silence added to the end of each stimulus. These periods of silence – in addition to an approximate

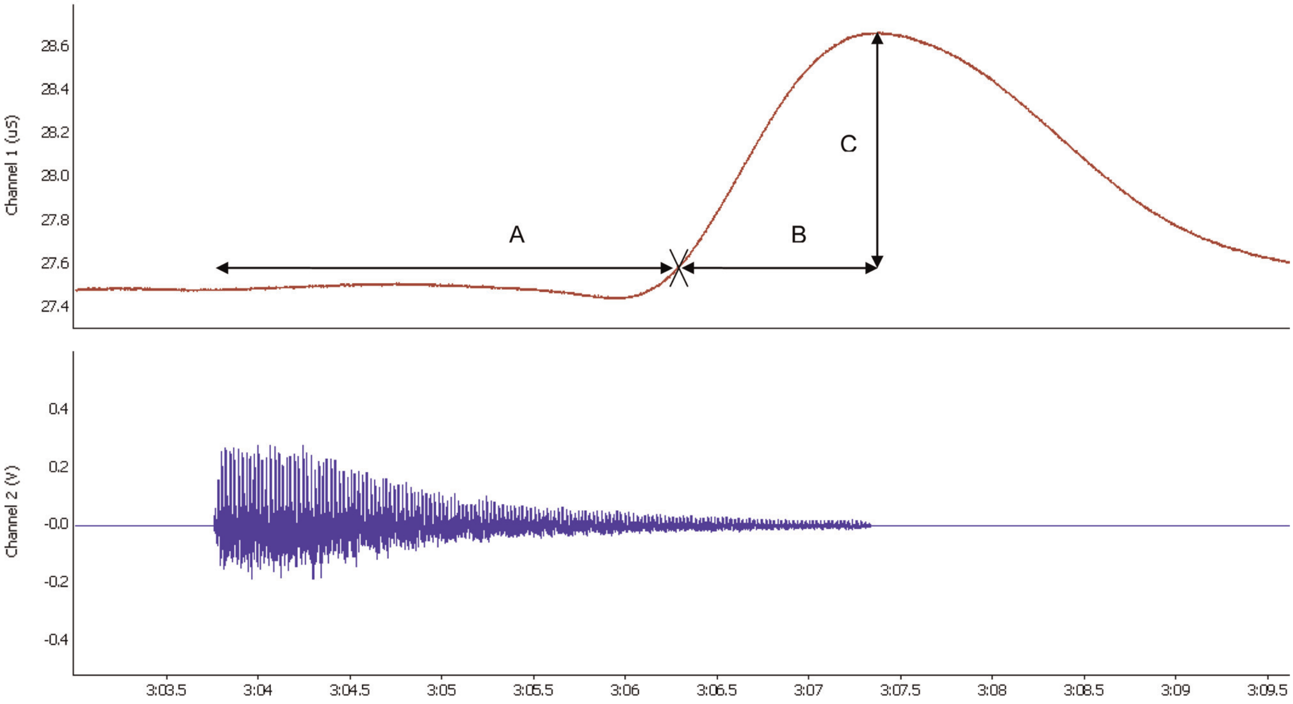

Figure 1. Example of an SCR recorded from a participant in response to a 3.6 s down-ramp stimulus. The *x*-axis represents the time elapsed during the experiment (min:sec). On the *y*-axis Channel 1 plots the SCL in μS units and Channel 2 shows stimulus input in volts. A = SCR latency; B = SCR rise time; C = SCR amplitude; X = approximate onset of SCR (0.05 μS increase relative to μS level at stimulus onset).

response time of 3 s for the computer-based loudness task – resulted in a mean intertrial interval of 15 s (range = 14–16 s). Pilot data indicated that up to 15 s ISI would suffice as an adequate time-frame for SCR to return to baseline levels before subsequent stimulus presentations.

Equipment

A LogicPro (Version 7.2.3) EXS24 sample playback program generated the violin stimulus and the Audacity (Version 1.3.3) sound editing program was used to create the 10 ms onset/offset ramps. Intensity levels for all stimuli were measured with a Brüel and Kjær Artificial Ear 4152 attached to a Brüel and Kjær Hand-held Analyzer 2250 using Sound Level Meter Software BZ-7222A. A PowerLab SCR amplifier with direct output to the main PowerLab 16/30 amplifier was used to record skin conductance at a sampling rate of 1,000 Hz and a range between 0.1 and 50 μS. The SCR amplifier uses a very low-current constant-voltage AC excitation (22 $\mathrm{mV_{rms}}$ at 75 Hz) to measure skin conductance using two dry bipolar electrodes that do not require special electrolytes. The AC current flows into the very low impedance input of a transimpedance amplifier which converts current into voltage. The resulting signal then passes through a synchronous rectifier to obtain DC voltage proportional to skin conductance, then via a modulator to produce a 400 Hz AC signal suitable for passing across the isolation barrier (which provides electrical protection for the participant). On the other side of the isolation barrier, the AC signal is multiplied then synchronously rectified, restoring a DC voltage again proportional to skin conductance. To reduce noise in the rectified signal, it is passed through a 1 Hz, second-order low-pass filter. This leaves the general signal trends unchanged but removes higher-frequency fluctuations. The PowerLab amplifier was connected by USB port to a Dell Optiplex GX270 personal computer (with a 3 GHz Intel Pentium 4 processor and a Microsoft Windows XP, Professional Version 2002, Service Pack 2 operating system) which ran the associated Chart (Version 5.3) software program.

The generation and presentation of the computer-based Visual Analogue Scale (VAS) response system, sound randomization, and protocol used for the explicit ratings of arousal and valence, in addition to the loudness perception task, were sequenced using the Music Experiment Development System (MEDS) (Kendall, 2000). Stimuli were presented binaurally through Sennheiser HD 25 headphones. The experiment was conducted in a sound-attenuated booth.

Procedure

Ethical approval was received prior to the commencement of the experiment. Participants first read an experiment information sheet, gave written informed consent, and received standardized instructions regarding the task. The skin conductance electrodes were then secured on the medial phalanges of participants' index and fourth fingers on their nondominant hand. After equipment was secured, a text version of instructions was presented on the experiment computer. Participants were instructed to listen to each sound in a trial and rate the perceived magnitude of loudness change on a computer-based VAS, ranging from "No-Change" to "Large Change," with a "Moderate-Change" in loudness as the midpoint of the scale. The order of the bipolar anchors on the VAS was reversed for every other participant to distribute any response bias toward a particular end of the scale. Skin conductance was recorded throughout the experiment while participants completed the loudness perception task. Each of the four stimuli was presented in a pseudorandom sequence, in two separate but continuous blocks: the first four and the latter four. This ensured the average serial position of each stimulus was controlled across the course of the experiment. Eight trials in total were presented to each participant. After this, the skin conductance transducers were removed and the participant completed the emotional arousal and valence rating tasks.

The emotional arousal rating task and valence rating task both consisted of four experimental stimuli representing all combinations of the two independent variables. Each trial consisted of one stimulus. In the emotional arousal task, participants were asked to indicate on a revised VAS the level of arousal they experienced from the sound, from "Calming" to "Arousing." In the valence task, participants were asked to indicate the level of pleasantness experienced from the sound on a scale from "Unpleasant" to "Pleasant." The midpoint of each scale represented a "Neutral" response (neither calming nor arousing, or neither unpleasant nor pleasant, respectively). The presentation order of these two tasks was counterbalanced across the experiment and the order of the bipolar anchors on each arousal and valence VAS was reversed for every other participant. The experiment took approximately 30 min.

Data Analysis

Data analysis for ratings of arousal, valence, and loudness change, and SCR magnitude was carried out in SPSS (Version 17) using separate within-subjects ANOVA. For SCR amplitude, latency, and rise time dependent measures, linear mixed-effects models (LME) were implemented using the lmer program (lme4 package; Bates & Sarkar, 2006) in the R software platform (Version 2.14). The LME approach was chosen to analyze SCR amplitude, latency, and rise time because of missing data resulting from trials that did not reach the 0.05 μS threshold (see Results section for SCR probability details). A traditional within-subjects ANOVA excludes each participant's entire data set if data are missing from any condition, leading to a considerable loss of statistical power. LME is capable of analyzing all data from all participants in a within-subjects design with missing cells. Consequently, using a within-subjects LME with missing data results in less loss of statistical power than a within-subjects ANOVA (Baayen, 2008; Pinheiro & Bates, 2000; Quené & Van den Bergh, 2004).

In the current LME analysis, the two levels of Intensity and Duration independent variables were designated fixed-effects, and random intercepts were determined per participant for each model fitting SCR amplitude, latency, and rise time data. The standard LME regression coefficients (*b*), standard errors (*SE*), and *t*-scores are reported for SCR amplitude, latency, and rise time main effects and interactions. In LME, an upper-bound degrees of freedom denominator used for *t*-values is calculated by the number of observations minus the number of fixed-effects (four in the current 2 × 2 design) (Baayen, Davidson, & Bates, 2008; Kliegl, Risse, & Laubrock, 2007). Consequently, *p*-values tend to be anti-conservative and increase the likelihood of Type I errors. To gain more conservative *p*-value calculations, the recommended Markov Chain Monte Carlo (MCMC) sampling method was implemented (Andrieu, De Freitas, Doucet, & Jordan, 2003; Baayen et al., 2008; Gelman & Hill, 2007) using the pvals.fnc() function in the LanguageR package (Baayen, 2011) with default specifications ($n = 10{,}000$ samples). Both original and MCMC procedures for calculating *p*-values in LME provided equivalent significance results ($\alpha = .05$).

Results

Judged Loudness Change

Loudness change was measured on a VAS; for each condition, a score of zero represents "No-Change" in loudness, whereas a score of 50 represents a "Large-Change" in loudness. A score of 25 represents a "Moderate-Change." First, there was a significant main effect Intensity, $F(1, 44) = 62.62$, $p < .001$, $\eta_p^2 = .59$. As can be seen in Figure 2, loudness change was significantly greater for up-ramps ($M = 41.32$, $SD = 6.23$), relative to down-ramps ($M = 30.88$, $SD = 10.49$). Second, there was a significant main effect of Duration, $F(1, 44) = 26.89$, $p < .001$, $\eta_p^2 = .38$. Loudness change was significantly greater for 3.6 s stimuli ($M = 38.38$, $SD = 9.70$), relative to 1.8 s stimuli ($M = 33.82$, $SD = 9.96$). The Intensity × Duration interaction was not significant, $F(1, 44) = 1.13$, $p > .05$, $\eta_p^2 = .03$. Overall the first hypothesis was supported: Loudness change was significantly greater for up-ramps relative to down-ramps, even though the physical intensity change within each up-ramp and down-ramp stimulus was identical (30 dB SPL).

Ratings of Emotional Arousal and Valence

Results of the subjective appraisal of the two dimensions of Russell's (1980) circumplex arousal-valence model of emotion are presented in Figure 3. Recall for the arousal VAS (Figure 3A), zero represents a "Calming" response and 50 represents an "Arousing" response, with a score of

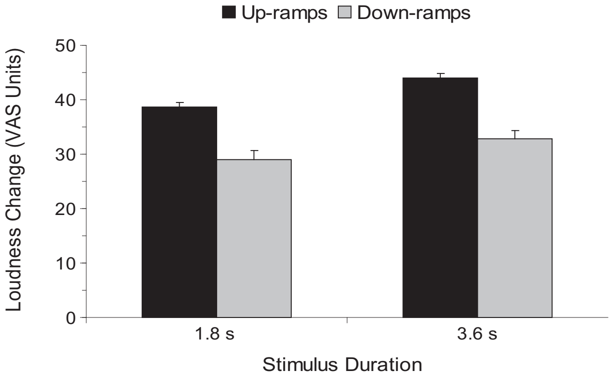

Figure 2. Mean ratings of loudness change for up-ramps and down-ramps as a function of stimulus duration. On the scale, zero represents "No-Change" in loudness and 50 represents a "Large-Change" in loudness. A score of 25 represents a "Moderate-Change" in loudness. Error bars represent standard error of the mean.

25 corresponding to a "Neutral" response (neither calming nor arousing). For the valence VAS (Figure 3B), zero represents an "Unpleasant" response and 50 represents a "Pleasant" response, with a score of 25 corresponding to a "Neutral" response (neither unpleasant nor pleasant).

For the arousal conditions, there was a significant main effect of Intensity, $F(1, 44) = 33.12$, $p < .001$, $\eta_p^2 = .43$. In support of the second hypothesis, up-ramps ($M = 38.33$, $SD = 7.22$) were rated significantly more arousing than down-ramps ($M = 28.90$, $SD = 11.01$). There was no significant main effect of Duration or significant Intensity × Duration interaction (F-values < .12). For the valence conditions, there was a significant main effect of Duration, $F(1, 44) = 4.12$, $p < .05$, $\eta_p^2 = .09$. The longer duration 3.6 s stimuli ($M = 16.36$, $SD = 10.68$) were rated significantly more unpleasant than the shorter 1.8 s stimuli ($M = 18.52$, $SD = 9.56$). There was no significant main effect of Intensity for valence ratings, nor was there a significant Intensity × Duration interaction (F-values < 2.16).

Event-Related SCR

First, SCR habituation over the course of the experiment was assessed by calculating the response probability (using the 0.05 μS threshold) from each specific trial placement, regardless of condition. As can be seen in Figure 4, a serial order effect is evident: the probability of SCRs declined as experiment trials increased. Specifically, the response probability of trials 5–8 was 27.78%, whereas the response probability of the first four trials was 67.78%. Therefore, only SCRs recorded in trials 1–4 were included in the analysis. Second, the proportion of SCRs elicited by up-ramps and down-ramps was analyzed. The proportion of responses attributed to up-ramp stimuli was 44.26% and the proportion of responses attributed to down-ramp stimuli was 55.74%. Frequency of occurrence of SCRs, comparing up- and down-ramps, was not significantly different, $\chi^2(1, N = 122) = 1.61$, $p > .05$. Finally, SCR magnitude, amplitude, latency, and rise time were log-transformed to normalize positive skew in the data (Tabachnick & Fidell, 2007).

Overall, electrodermal results show that relative to down-ramps, up-ramps elicited significantly smaller SCR magnitude in conjunction with longer SCR rise times that approached significance ($\alpha = .05$). SCR latency and amplitude were equivalent between the two directions of intensity change. Statistics for each SCR-dependent measure will now be presented.

SCR Magnitude

For mean SCR magnitude, there was a significant main effect of Intensity, $F(1, 44) = 6.70$, $p < .05$, $\eta_p^2 = .13$. As can be seen in Figure 5A, SCR magnitude was significantly greater in response to down-ramps ($M = .396$, $SD = 0.084$) relative to up-ramps ($M = .356$, $SD = 0.120$). There was no

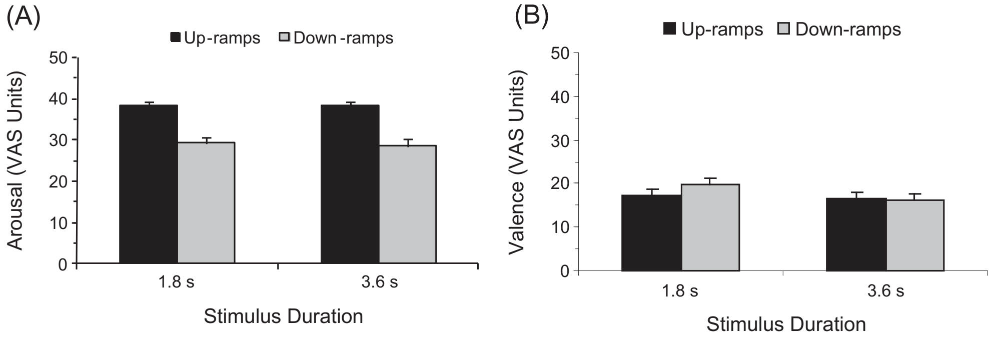

Figure 3. Mean arousal and valence ratings for up-ramps and down-ramps as a function of stimulus duration. For the arousal scale (A), zero represents "Calming" and 50 represents "Arousing," with a score of 25 corresponding to a "Neutral" response. For the valence scale (B), zero represents "Unpleasant" and 50 represents "Pleasant," with a score of 25 corresponding to a "Neutral" response. Error bars represent standard error of the mean.

significant main effect of Duration or Intensity × Duration interaction (F-values < .92).

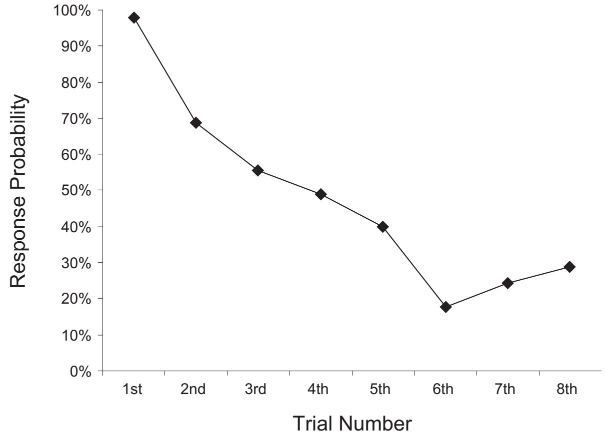

Figure 4. SCR probability from 45 participants as a function of trial number, collapsed across variables.

SCR Amplitude

For mean SCR amplitude, there was no significant main effect of Intensity, $b = .014$, $SE = .035$, $t = .385$, $p > .05$, or Duration, $b = -.031$, $SE = .036$, $t = -.871$, $p > .05$, and no significant Intensity × Duration interaction, $b = .022$, $SE = .036$, $t = .609$, $p > .05$ (see Figure 5B).

SCR Latency

For mean SCR latency, there was no significant main effect of Intensity, $b = .012$, $SE = .014$, $t = .874$, $p > .05$, or Duration, $b = .001$, $SE = .014$, $t = .072$, $p > .05$, and no significant Intensity × Duration interaction, $b = -.021$, $SE = .014$, $t = -1.488$, $p > .05$ (see Figure 5C).

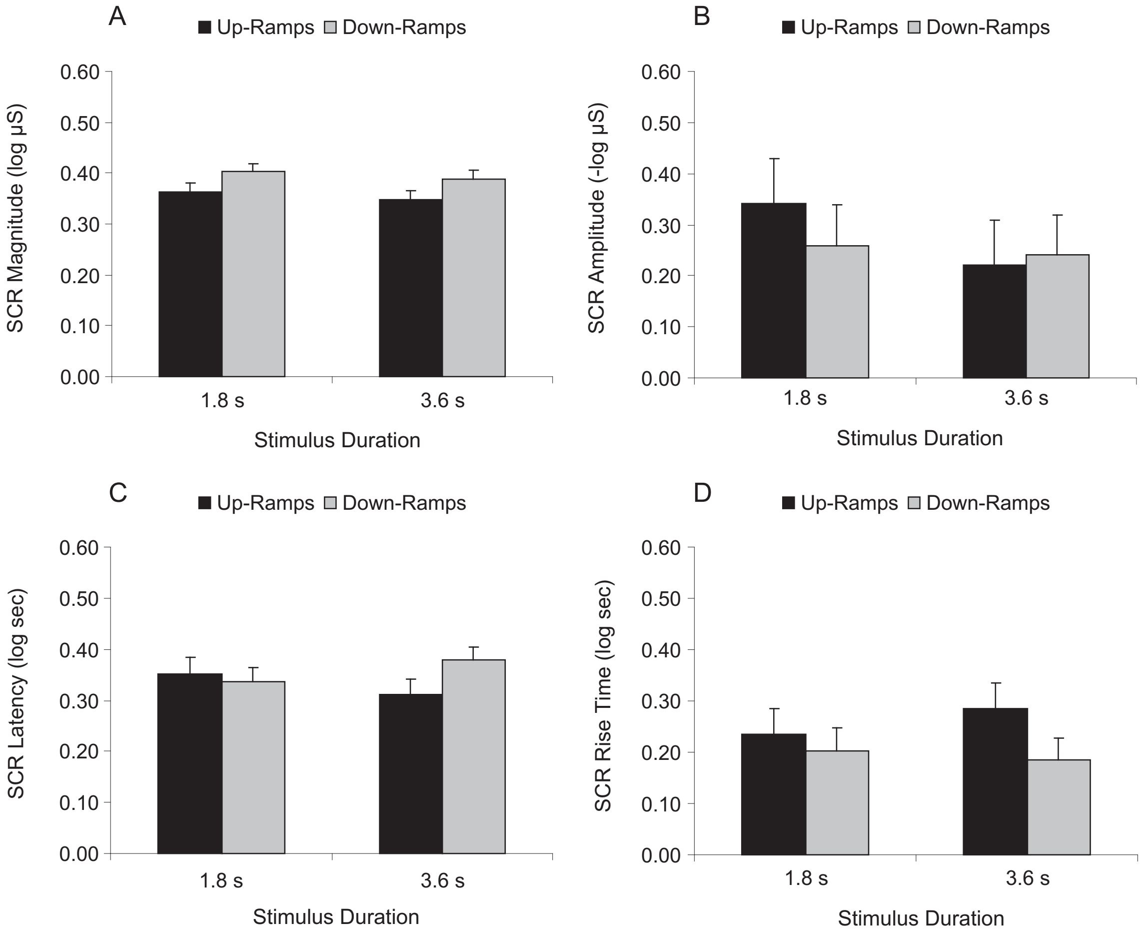

Figure 5. Results of each aspect of the SCR as a function of stimulus duration: (A) SCR magnitude refers to mean SCL between 4 and 5 s after stimulus onset, corrected for SCL 1 s baseline before stimulus onset; (B) SCR amplitude refers to the mean microSiemens (µS) increase in SCR, relative to baseline SCL at stimulus onset; (C) SCR latency refers to the mean duration from stimulus onset to response onset; and (D) SCR rise time refers to the mean duration from response onset to response peak. Error bars represent standard error of the mean.

SCR Rise Time

For mean SCR rise time, the main effect of Intensity approached significance, $b = -.032$, $SE = .018$, $t = -1.790$, $p = .076$. SCR rise time (duration of the SCR) was marginally longer for up-ramps ($M = .259$, $SD = 0.237$) relative to down-ramps ($M = .193$, $SD = 0.271$). There was no significant main effect of Duration, $b = -.013$, $SE = .018$, $t = -.721$, $p > .05$, and no significant Intensity × Duration interaction, $b = .016$, $SE = .019$, $t = .842$, $p > .05$ (see Figure 5D).

Discussion

An experiment investigated the overarching hypothesis that ratings of emotional arousal and valence in response to musical up-ramps are associated with judged loudness change and a characteristic psychophysiological response. Relative to musical down-ramps matched on acoustic attributes such as the frequency, average intensity, and range of intensity change, up-ramps elicited significantly higher ratings of emotional arousal and loudness change. However, a significantly greater magnitude of psychophysiological increase was observed in response to musical down-ramps. We now consider each specific hypothesis in turn.

Loudness Change, Emotional Arousal, and Valence

We first hypothesized an overestimation of loudness change in response to up-ramps, relative to down-ramps. This hypothesis was supported. Musical up-ramp stimuli were judged to change significantly more in loudness, even though the intensity change within each up-ramp and down-ramp was identical (30 dB SPL). This result replicates a range of studies (Bach et al., 2009; Neuhoff, 1998, 2001; Olsen et al., 2010; Seifritz et al., 2002) that report a greater perceived magnitude of loudness change than is physically present in up-ramp stimuli relative to down-ramps. However, there is evidence that poststimulus judgments of loudness change are heavily influenced by the end-level of the stimulus, and not the entire intensity change of the stimulus (Olsen & Stevens, 2010; Olsen et al., 2010; Susini et al., 2007; Susini, Meunier, Trapeau, & Chatron, 2010; Teghtsoonian, Teghtsoonian, & Canévet, 2005). In the present study, up-ramps ended on 90 dB SPL, whereas down-ramps ended on 60 dB SPL. The 30 dB SPL difference in end-level between the two directions of change may have biased direct judgments of loudness change. Therefore, caution is needed when assessing loudness change data where up-ramp and down-ramp offset levels are not balanced.

Second, we hypothesized higher ratings of emotional arousal and negative valence (unpleasantness) in response to musical up-ramps, relative to down-ramps. This hypothesis was partially supported. Explicit ratings of arousal show that regardless of stimulus duration, up-ramps were rated significantly more arousing than down-ramps. This result replicates subjective arousal results reported in Bach et al. (2009). However, ratings of valence were not intensity-direction specific. Participants rated the 3.6 s stimuli more unpleasant than the 1.8 s stimuli, regardless of the intensity change profile. Relatively high ratings of unpleasantness in response to dynamic pure tones have previously been associated with an increase of stimulus duration, from 1 s to 3 s (Tajadura-Jiménez, Väljamäe, Asutay, & Västfjäll, 2010). Tajadura-Jiménez et al. suggest that the longer stimulus duration affords listeners time to properly evaluate the sound and prepare for response. In the present study, the spectral content and/or timbre of the four-note dissonant diminished chord is an additional factor in significantly greater ratings of unpleasantness in response to 3.6 s stimuli, relative to 1.8 s stimuli. The dissonant diminished chord is a relatively "harsh" sound that is often associated with a heightened sense of arousal, unpleasantness, and tension in musical and cinematic contexts (Huron, 2006; Juslin & Västfjäll, 2008; Olsen et al., 2010). All stimuli were likely to be perceived by participants as unpleasant to begin with; the extended duration merely exacerbated the unpleasantness.

Effects of Intensity Change on SCR

Analysis of the psychophysiological response to dynamic acoustic intensity was made using an event-related SCR paradigm. Up-ramps and down-ramps that change over a 30 dB SPL range (60–90 dB SPL and 90–60 dB SPL, respectively) elicit a statistically equivalent number of SCRs characterized by marginal differences in response rise times, and no differences in response amplitude and latency. It was initially predicted that up-ramps would elicit faster SCR onset (as evident by shorter SCR latencies) because continuous increases of intensity demand behavioral priority. This is arguably due to the relationship between intensity and looming auditory motion in the environment (Neuhoff, 1998, 2001). Equivalent response latency when measured from stimulus onset to SCR onset shows that the timing of SCR onset is primarily affected by stimulus onset and not the dynamic intensity contour that follows. The temporal characteristics of psychophysiological response onset to relatively short up-ramps do not follow the predicted pattern and conflict with previously reported behavioral data. That is, reaction times in response to up-ramps are reportedly faster than equivalent down-ramps (Bach et al., 2009; Bach et al., 2008; Tajadura-Jiménez et al., 2010). From the present data, there is no parallel psychophysiological "behavioral priority" in the sense of faster SCR to continuous increases of intensity.

Equivalent SCR amplitude between up-ramps and down-ramps shows that once the 0.05 μS threshold for an SCR is met, the amplitude of the SCR increases to a maximal point that does not systematically vary with intensity direction or stimulus duration. Full-motion cues from virtual moving sounds – and not just intensity change alone – are necessary to achieve significant differences in SCR amplitude (Bach et al., 2009). However, SCR duration (rise time) is affected by the direction of acoustic intensity change. Relative to down-ramps, up-ramps elicit marginally longer SCRs, regardless of stimulus duration. A sustained psychophysiological response to increasing intensity and a looming (approaching) object in the environment would lead to a heightened state of arousal associated with a rapid response to a potentially threatening event (Bach et al., 2008; Hall & Moore, 2003; Maier & Ghazanfar, 2007; Seifritz et al., 2002). Indeed, explicit ratings of arousal show that regardless of stimulus duration, up-ramps were rated significantly more arousing than down-ramps. This is consistent across studies (Bach et al., 2009; Tajadura-Jiménez et al., 2010). As the SCR amplitude was equivalent between up-ramps and down-ramps, it is likely that higher ratings of emotional arousal in response to up-ramps are more closely associated with the duration (rise time) of psychophysiological increase, and not just amplitude per se. This result sheds greater light on the relationship between implicit indices of psychophysiological arousal and explicit judgments of psychological arousal in response to dynamic acoustic intensity.

Finally, the magnitude of psychophysiological response was significantly greater for musical down-ramps, relative to up-ramps. SCR magnitude includes averaged data across all trials, regardless of whether skin conductance level increased beyond the 0.05 μS threshold. Greater SCR magnitude in response to musical down-ramps contrasts Experiment 2 in Bach et al. (2009), where there was greater SCR magnitude to 2 s square-wave up-ramps using equivalent response criteria. Differences in SCR magnitude between the present study and Bach et al. (2009) can be reconciled if we consider the onset characteristics of the down-ramps. As much of the psychophysiological response to these short dynamic stimuli is weighted on stimulus onset, it is plausible that the 90 dB SPL onset of a down-ramp in the present study may have, on average, increased overall SCL on trials that did not reach the 0.05 μS threshold. Relative to the 60 dB SPL up-ramp onset, the 90 dB SPL onset would have exacerbated SCR magnitude to down-ramps where all trials (including no responses) are calculated. In Bach et al. (2009), dynamic intensity sweeps were comparatively smaller (20 dB SPL as opposed to 30 dB SPL here) and their 85 dB SPL down-ramp onset may not have elicited such a pronounced effect. The influence of dynamic intensity onset on psychophysiological arousal can be investigated by systematically varying the regions of intensity change; for example, incorporating low (50–70 dB) and high (70–90 dB) regions of intensity change used in previous psychoacoustic experiments (e.g., Olsen et al., 2010; Teghtsoonian et al., 2005).

It is important to note here that a control condition would assist in assessing the influence of spectral and timbral differences in stimuli across psychophysiological investigations (cf. Bach et al., 2008, 2009). Future work that systematically manipulates stimuli from recent auditory looming/dynamic intensity research, such as pure tones, square-waves, white-noise, and more ecologically valid stimuli such as musical timbres and simple speech (e.g., a vowel), will shed light on this issue.

Implications for Perception of Music and Auditory Warnings

The rate of intensity change within up-ramp and down-ramp conditions in the present study did not have a significant impact on SCR. That is, there were no differences in SCR amplitude or magnitude between the faster rate of intensity change inherent in 1.8 s stimuli (16.67 dB SPL change per second) relative to 3.6 s stimuli (8.33 dB SPL change per second). This result is surprising, as increased self-reported emotional arousal and shivers down the spine have been correlated with fast rates of intensity change in Western Classical musical excerpts (Schubert & Dunsmuir, 1999; Yasuda, 2009). The difference may lie in the relatively short duration and reduced musical complexity of the present set of stimuli. Nevertheless, future studies could investigate the relationship between rate of acoustic intensity change, psychophysiological and self-report measures, in conjunction with computational analyses of *dynamic* psychoacoustic features of increasingly complex music that correlate with emotional response (e.g., Dean, Bailes, & Schubert, 2011). Embedding systematic intensity change in a more complex and temporally extended musical context may also reduce the relatively fast psychophysiological habituation observed in the present study.

Furthermore, the psychophysiological response to acoustic intensity change within musical stimuli can be used to investigate cognitive-emotional appraisal of crescendo and diminuendo. For example, musicological analyses suggest that a long and gradual crescendo (relative to a short, abrupt diminuendo) is used by composers to maintain listeners' attention (Huron, 1992). An extended crescendo is also likely to lead to a longer psychophysiological arousal response that may intensify the perceived or felt emotion the listener experiences during that portion of a musical piece. Importantly, we emphasize that stimulus *change* is the crucial and hitherto overlooked feature in music and emotion research.

Finally, the present study adds support to the theory of a rapid human response to a rapid intensity increase that is characteristic of the intrinsic warning properties of looming auditory motion. In another applied context, a set of "biased" responses to looming auditory motion are highly relevant to the design of auditory warnings that are used in complex operational environments (Edworthy & Hellier, 2006; Keller & Stevens, 2004; Stephan, Smith, Martin, Parker, & McAnally, 2006). For example, critical information should be recognized with greater speed and accuracy if carried within an up-ramp or following an up-ramp, relative to steady-state presentation. In a monitoring context, change should be most easily detected when signaled by an up-ramp. From a design perspective, rapid up-ramps have the advantage of exploiting psychophysiological response mechanisms without adding excessive, sustained intensities to environments that may already be noisy or auditorily crowded.

Conclusion

A clearer picture of human perception in response to acoustic intensity dynamics is now emerging from complementary psychophysiological and psychophysical reports. Relative to down-ramps, up-ramps are perceived louder, longer in duration, and cover a greater range of loudness change (e.g., Grassi & Darwin, 2006; Neuhoff, 1998; Olsen et al., 2010; Susini et al., 2007). Differences in sensory processes such as neural persistence or adaptation at peripheral and/or central stages of auditory processing are potential neural mechanisms (DiGiovanni & Schlauch, 2007; Ries et al., 2008). A characteristic psychophysiological response is associated with perceptual differences between up-ramp and down-ramp stimuli, whereby up-ramps elicit marginally longer increases of autonomic arousal, in conjunction with higher ratings of emotional arousal. In addition to automatic subcortical activation in response to acoustic intensity change in early auditory processing (Juslin & Västfjäll, 2008), persistence of neural activation and psychophysiological arousal are likely associated with the "salience" of up-ramp stimuli and their contribution to music-induced emotion. The likely mechanism is a neural network (Bach et al., 2008; Hall & Moore, 2003; Seifritz et al., 2002) that responds to real and apparent looming auditory motion in a range of listening contexts.

Acknowledgment

This research was supported by an Australian Postgraduate Award held by the first author and an Australian Research Council Discovery Project grant (DP0771890) held by the second author. We thank Jacques Launay for statistical advice and members of the MARCS Institute Music Cognition and Action Group for helpful comments on an earlier draft.

References

Andreassi, J. L. (2000). *Psychophysiology: Human behavior and physiological response* (4th ed.). Mahwah, NJ: Erlbaum.

Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. I. (2003). An introduction to MCMC for machine learning. *Machine Learning, 50*, 5–43.

Baayen, R. H. (2008). *Practical data analysis for the language sciences with R*. Cambridge, MA: Cambridge University Press.

Baayen, R. H. (2011). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics" (Version 1.2) [Computer Software]. Vienna, Austria: R Foundation for Statistical Computing.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bach, D., Neuhoff, J., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology, 74*, 28–33.

Bach, D., Schachinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., . . . Seifritz, E. (2008). Rising sound intensity: An intrinsic warning cue activating the amygdala. *Cerebral Cortex, 18*, 145–150.

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*, 43–64.

Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. I. E. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research, 46*, 337–349.

Bates, D., & Sarkar, D. (2006). lme4: Linear mixed-effect models using S4 classes R package (Version 2.14) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition, 80*, 1–10.

Dean, R. T., & Bailes, F. (2010). A rise-fall temporal asymmetry of intensity in composed and improvised electroacoustic music. *Organised Sound, 15*, 147–158.

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.

DiGiovanni, J. J., & Schlauch, R. S. (2007). Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds. *Ecological Psychology, 19*, 239–264.

Edworthy, J., & Hellier, E. (2006). Alarms and human behaviour: Implications for medical alarms. *British Journal of Anaesthesia, 97*, 12–17.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 223–248). Oxford, UK: Oxford University Press.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition & Emotion, 17*, 25–40.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Grassi, M., & Darwin, C. J. (2006). The subjective duration of ramped and damped sounds. *Perception and Psychophysics, 68*, 1382–1392.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception, 24*, 297–314.

Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and musico-acoustic correlates of the chill response. *Music Perception, 24*, 473–484.

Hall, D. A., & Moore, D. R. (2003). Auditory neuroscience: The salience of looming sounds. *Current Biology, 13*, R91–R93.

Heinlein, C. P. (1928). The affective characters of the major and minor modes in music. *Journal of Comparative Psychology, 8*, 101–142.

Huron, D. (1990). Crescendo/diminuendo asymmetries in Beethoven's piano sonatas. *Music Perception, 7*, 395–402.

Huron, D. (1991). The ramp archetype: A score-based study of musical dynamics in 14 piano composers. *Psychology of Music, 19*, 33–45.

Huron, D. (1992). The ramp archetype and the maintenance of passive auditory attention. *Music Perception, 10*, 83–91.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception, 23*, 319–330.

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

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559–575.

Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied, 10*, 3–12.

Kendall, R. A. (2000). Music Experiment Development System (MEDS) [Computer Software]. Retrieved from http://www.schoolofmusic.ucla.edu/pages/Kendall/meds.htm

Khalfa, S., Peretz, I., Blondin, J.-P., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters, 328*, 145–149.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n + 2. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1250–1255.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51*, 336–353.

Lang, P., Bradley, M., & Cuthbert, B. (1998). Emotion and motivation: Measuring affective perception. *Journal of Clinical Neurophysiology, 15*, 397–408.

Maier, J. X., & Ghazanfar, A. A. (2007). Looming biases in monkey auditory cortex. *Journal of Neuroscience, 27*, 4093–4100.

Nagel, F., Kopiez, R., Grewe, O., & Altenmüller, E. (2008). Psychoacoustical correlates of musically induced chills. *Musicae Scientiae, 12*, 101–113.

Nanez, J. (1988). Perception of impending collision in 3- to 6-week-old human infants. *Infant Behavior and Development, 11*, 447–463.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature, 395*, 123–124.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology, 13*, 87–110.

Neuhoff, J. G. (2004). *Ecological Psychoacoustics*. London, UK: Elsevier Academic Press.

Olsen, K. N., & Stevens, C. J. (2010). Perceptual overestimation of rising intensity: Is stimulus continuity necessary? *Perception, 39*, 695–704.

Olsen, K. N., Stevens, C. J., & Tardieu, J. (2010). Loudness change in response to dynamic acoustic intensity. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1631–1644.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication, 43*, 103–121.

Ries, D. T., Schlauch, R. S., & DiGiovanni, J. J. (2008). The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds. *Journal of the Acoustical Society of America, 124*, 3772–3783.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 303–316.

Schlauch, R. S., Ries, D. T., & DiGiovanni, J. J. (2001). Duration discrimination and subjective duration for ramped and damped sounds. *Journal of the Acoustical Society of America, 109*, 2880–2887.

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception, 21*, 561–585.

Schubert, E., & Dunsmuir, W. (1999). Regression modelling continuous data in music psychology. In Suk Won Yi (Ed.), *Music, mind, and science* (pp. 298–352). Seoul, South Korea: National University Press.

Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schachinger, H., . . . Di Salle, F. (2002). Neural processing of auditory looming in the human brain. *Current Biology, 12*, 2147–2151.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music, 19*, 110–120.

Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *Journal of the Acoustical Society of America, 107*, 3358–3368.

Stephan, K. L., Smith, S. E., Martin, R. L., Parker, S. P. A., & McAnally, K. I. (2006). Learning and retention of associations between auditory icons and denotative referents: Implications for the design of auditory warnings. *Human Factors, 48*, 279–287.

Susini, P., McAdams, S., & Smith, B. K. (2007). Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings. *Acta Acustica united with Acustica, 93*, 623–631.

Susini, P., Meunier, S., Trapeau, R., & Chatron, J. (2010). End level bias on direct loudness ratings of increasing sounds. *Journal of the Acoustical Society of America, 128*, EL163–EL168.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MD: Allyn & Bacon.

Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: The emotional impact of approaching and receding sound sources. *Emotion, 10*, 216–229.

Teghtsoonian, R., Teghtsoonian, M., & Canévet, G. (2005). Sweep-induced acceleration in loudness change and the "bias for rising intensities". *Perception and Psychophysics, 67*, 699–711.

Venables, P., & Christie, M. (1980). Electrodermal activity. In I. Martin & P. Venables (Eds.), *Techniques in psychophysiology* (pp. 3–67). Chichester, UK: Wiley.

Yasuda, S. (2009). A psychological study of strong experiences induced by listening to music: Relationship between subjectively evaluated physical reactions and change in volume while listening. In C. Stevens, E. Schubert, B. Kruithof, K. Buckley, & S. Fazio (Eds.), *Proceedings of the 2nd International Conference on Music Communication Science (ICoMCS2)*. Sydney: HCSNet, University of Western Sydney.

Accepted for publication: August 24, 2012

Kirk N. Olsen

University of Western Sydney, Australia
Locked Bag 1797
Penrith, NSW 2751
Tel. +61 2 9772 6660
Fax +61 2 9772 6326
E-mail k.olsen@uws.edu.au