Psychomusicology: Music, Mind, and Brain
2014, Vol. 24, No. 2, 147–156

0275-3987/14/$12.00 http://dx.doi.org/10.1037/pmu0000044

A Continuous Measure of Musical Engagement Contributes to Prediction of Perceived Arousal and Valence

Kirk N. Olsen, Roger T. Dean, and Catherine J. Stevens
University of Western Sydney

A listener's propensity to perceive affect as expressed by music can arise from factors such as acoustic features and culturally learned expectations. Studies investigating the link between musical flow and perceived affective content by means of continuous response measures and a 2-dimensional circumplex framework of affect (i.e., arousal and valence) have given positive results. For example, time series models of perceived arousal in response to Western classical and electroacoustic music reveal a significant predictive influence of acoustic parameters such as intensity and spectral flatness. Acoustic parameters generally provide weaker models of perceived valence. Here we test the hypothesis that a continuous measure of musical engagement can be a significant predictor of perceived arousal and perceived valence, and will enhance time series models of affect based on acoustic parameters alone. Thirty-five nonmusicians continuously rated their level of engagement while listening to 5 Western classical and electroacoustic music excerpts. Grand unweighted mean engagement time series for each piece from all 35 participants were used to model continuous-response time series of perceived arousal and perceived valence. The hypothesis was partially supported: in univariate autoregressive analyses, 1 of the valence and 2 of the arousal models were strongly improved by adding engagement as a predictor; and in a further 2 of each, engagement made a minor contribution. In the remaining 2 models of valence and 1 of arousal, engagement was not pertinent. In multivariate (vector autoregressive) models, relating simultaneously both arousal and valence to acoustic parameters, engagement had a role in every case. It is concluded that listener engagement can play a mediating role in perceived affective response to music.

Keywords: affect, engagement, music, perception, time series analysis

Listeners' affective response to music has received extensive empirical and theoretical treatment that considers culturally specific factors such as tonal familiarity, and surface structures such as acoustic intensity (Balkwill & Thompson, 1999; Balkwill, Thompson, & Matsunaga, 2004; Dean, Bailes, & Schubert, 2011; Grewe, Nagel, Kopiez, & Altenmüller, 2007; Juslin, 2013; Juslin & Sloboda, 2001; Juslin & Västfjäll, 2008; Meyer, 1956; Olsen & Stevens, 2013). The primary framework used to investigate the link between music and affect is the two-dimensional circumplex model (Russell, 1980). In this framework, the dimension of "arousal" is commonly characterized in terms of activation, with anchors such as active/aroused and passive/calm (Schubert, 2010). The second dimension of "valence" comprises positive and negative anchors and may be conceptualized as the "pleasantness" of a stimulus. Other multidimensional models have also been proposed that further divide arousal into "energy arousal" and "tension arousal" (e.g., Thayer, 1978, 1986).

Russell's (1980) circumplex model is a robust framework apt for research on perceived affect in visual and auditory domains (Bradley & Lang, 2000b; Eerola & Vuoskoski, 2011). For example, affective pictures and naturally occurring sounds are widely represented across both dimensions (Bradley & Lang, 2000a, 2000b). Studies using retrospective ratings of affect in response to music (i.e., measured after a listener has heard a musical excerpt) show that perceived arousal is significantly associated with acoustic cues such as intensity, spectral flux, and spectral entropy (Gabrielsson & Lindström, 2010; Gingras, Marin, & Fitch, 2013; Ilie & Thompson, 2006; Leman, Vermeulen, De Voogdt, Moelants, & Lesaffre, 2005). Furthermore, by using a two-dimensional emotion–space interface (Schubert, 1999), continuous real-time measurements of perceived affect recorded throughout a musical excerpt result in time series models that significantly predict perceived arousal from a small number of musical features (Bailes & Dean, 2012; Dean & Bailes, 2010; Schubert, 2004, 2013). For example, acoustic intensity profiles in Western classical and electroacoustic music can significantly predict continuous changes in perceived arousal, and this has been supported by causal experiments in which the intensity profiles of pieces have been manipulated (Dean & Bailes, 2011). Timbral features such as spectral centroid and spectral flatness have weaker effects (Dean & Bailes, 2010, 2011). On the other hand, perceived valence in response to music is commonly characterized by high variability in retrospective responses (e.g., Gomez & Danuser, 2004; Leman et

Kirk N. Olsen and Roger T. Dean, The MARCS Institute, University of Western Sydney; and Catherine J. Stevens, The MARCS Institute and School of Social Sciences and Psychology, University of Western Sydney.

This research was supported by an Australian Research Council Discovery Project grant (DP120102426) held by the second and third authors. We thank Annabel Cohen, two anonymous reviewers, and the MARCS Institute Music Cognition and Action research group for helpful comments on an earlier draft.

Correspondence concerning this article should be addressed to Kirk N. Olsen, The MARCS Institute, University of Western Sydney, Locked Bag 1797, Penrith, NSW 2751, Australia. E-mail: k.olsen@uws.edu.au. Web: http://marcs.uws.edu.au

al., 2005) and low predictive power from computational models (e.g., Bailes & Dean, 2012; Dean & Bailes, 2010; Korhonen, Clausi, & Jernigan, 2006; Schubert, 2004). Furthermore, causal links between acoustic aspects of music and perceived valence have yet to be demonstrated.

So why is perceived valence predicted less well by acoustic cues than perceived arousal? Valence may be closely associated with more personal factors that could be culture-specific and less obligatory than the effects observed from acoustic cues on perceived arousal; for example, the individual motivational aspects of attention and interest that underpin listeners' engagement with a piece of music (Broughton, Stevens, & Schubert, 2008; Geringer & Madsen, 2000–2001; Madsen & Geringer, 2008; Thompson, 2007). In general terms, engagement refers to an active, constructive, focused interaction with one's social and physical environment that relates to cognitive, behavioral, and importantly for the present study, affective elements of motivation (Broughton et al., 2008; Furrer & Skinner, 2003; Reeve, Jang, Carrell, Jeon, & Barch, 2004). In music, a listener's real-time engagement is likely to be associated with greater attention and interest, and correlate (either negatively or positively) with affective valence responses such as enjoyment and pleasantness (or lack thereof)[1].

As music is dynamic and unfolding through time, the present study was designed to use time series modeling to investigate whether a continuous measure of engagement throughout a piece of music serves as a significant predictor of continuous ratings of perceived arousal and perceived valence in particular. We also asked whether a continuous measure of engagement significantly enhances time series models of perceived arousal and perceived valence, based solely on continuously varying global acoustic features of intensity and spectral flatness in response to diverse musical genres (e.g., Western classical and electroacoustic music; see Bailes & Dean, 2012). Acoustic intensity is defined as a sound's power over unit area (most commonly measured in dB SPL). Spectral flatness is a global parameter of timbre and is measured as the ratio of the geometric mean to the arithmetic mean of the power spectrum, here expressed on a logarithmic scale. It is predicted that time series models of perceived arousal and perceived valence will be significantly strengthened when including continuous ratings of musical engagement, in addition to acoustic parameter profiles such as acoustic intensity and spectral flatness. To quantify this prediction, Bayesian information criterion (BIC) will be calculated as the basis for model selection and subsequently used to compare time series models of arousal and valence that do not include real-time engagement as a predictor. The relative influence of acoustic variables (intensity and spectral flatness) will also be assessed.

The present study uses a sample of musical excerpts from Western classical and electroacoustic music that include a range of instrumentation and sound sources with diverse complexity, familiarity, and stasis. For example, a piano concerto by Mozart represents a relatively familiar and complex musical example from the Western classical tonal tradition, with dynamic agency in the form of soloist entries. A piano piece composed by Webern serves as a single-timbre instrumental piece, but contrasted with Mozart by its pointillist and atonal quality and relative rhythmic monotony. Three electroacoustic excerpts by Dean, Wishart, and Xenakis incorporate varying degrees of noninstrumental sound sources and timbral complexity, with Wishart the most complex in its timbral variation and Dean and Xenakis showing very limited variation in timbre and intensity (see coefficients of variation in Results section, Table 3). The choice of this sample of excerpts was further motivated by the opportunity to use real-time engagement data to build on past models of perceived valence and perceived arousal that were derived from the experiment in Bailes and Dean (2012), which implemented an identical paradigm to the present study. Specifically, it was hypothesized that—

Hypothesis 1: A continuous measure of musical engagement is a predictor of perceived arousal and valence;

Hypothesis 2: Musical engagement is a strong predictor and may enhance Bailes and Dean's (2012) previous time series models of perceived arousal and particularly those of perceived valence.

Method

Participants

The sample consisted of 35 adult participants recruited from the University of Western Sydney (26 females and 9 males; M = 21.31 years, SD = 6.13, range = 18–44 years). All reported normal hearing. Participants had a median Ollen Musical Sophistication Index (OMSI) (Ollen, 2006) of 142 (range = 19–445), which qualifies them as "not musically sophisticated."

Stimuli, Materials, and Equipment

Five excerpts of music were presented in random order for each participant in the experiment (see Appendix for URLs containing examples of each piece of music):

1. Wolfgang Amadeus Mozart, *Piano Concerto 21, K467* (2'19"): Performed by Daniel Barenboim and the English Chamber Orchestra, recorded 1969, and digitally remastered in 1986 from HMV 5 86740 2. Opening of the 3rd Movement, Allegro Vivace.

2. Anton Webern (1937) *Piano Variations Op. 27* (3'11"): The second (sehr schnell) and third (ruhig fliessend) movements were excerpted from a performance by Glenn Gould (1964) in the film "The Alchemist." The excerpt begins at about 1'45. This piano work is included as an instrumental, pointillist contrast to the other four pieces.

3. Roger Dean (2003) *soundAFFECTS* (3'01"): The excerpt from this composition featured filtered noise and is part of an audiovisual work for performance and for the web (Brewster, Smith, & Dean, 2004). Only the audio portion was presented in this experiment.

[1] It is important here to distinguish between different forms of musical engagement. The present study focuses on listeners' real-time engagement with music during a listening experience (i.e., how engaging they find a piece of music throughout the continuous listening process), as opposed to listeners' active engagement with music, which refers to one's level of involvement in musical practice, performance, and so forth (Chin & Rickard, 2012).

4. Trevor Wishart (1977) *Red Bird, a political prisoner's dream* (3'16"): This was excerpted from a recording on UbuWeb of this 45-min piece for tape, which has a strong narrative.

5. Iannis Xenakis (1962) *Bohor* (3'15"). A four-track work for tape, from which a stereo recording was excerpted from EMF CD 003.

In addition to these five excerpts, two practice trials were presented: excerpts of the first movement from Mozart's *Symphony No. 40, K550* (1'18") and of Xenakis's *Orient-Occident* (1'29"). Each stimulus excerpt in the experiment was presented as an .aiff stereo 16 bit audio file with a 44.1 kHz sampling rate. Paper materials comprised—(a) the 10-question OMSI (Ollen, 2006) assessing an individual's level of musical sophistication, in addition to basic demographic information; (b) the Absorption in Music Scale (AIMS) (Sandstrom & Russo, 2013), a 34-item music-specific scale developed from the commonly used but very general Tellegen Absorption Scale (Tellegen & Atkinson, 1974); (c) a familiarity question asking participants to rate their familiarity with each piece using five options, from 0: *I have never heard anything like this before* to 5: *I often listen to this piece of music*, with *I have heard this piece* at the midpoint of the scale; and (d) a likability question, asking participants to rate on a 5-point scale how much they liked each piece of music, from 0: *Really dislike* to 5: *Really like*, with *Neither like nor dislike* at the midpoint of the scale. Table 1 reports descriptive statistics for ratings of familiarity and liking for each excerpt of music. Measures of absorption, likability, and familiarity were not significantly correlated with individuals' ratings of musical engagement (obtained as an average across a piece for the continuous engagement measure). Therefore these measures were not used further in quantitative modeling.

The experiment was conducted in a sound-attenuated booth, and stimuli were presented binaurally through Sennheiser HD25 headphones. An Apple MacBook Pro laptop computer (System 10.6.2) using a custom written Java application was responsible for displaying on-screen instructions, the horizontal engagement scale, presentation of stimuli, and continuous recording of data. The engagement scale ranged from "Not Engaged" on the far left of the scale to "Engaged" on the far right, with "Neutral" as the midpoint. Participants made their continuous engagement ratings by using a computer mouse to move the on-screen arrow cursor along the scale. Ratings were continuously recorded and averaged in the application to provide data at a sampling rate of 2 Hz.

Table 1
Descriptive Statistics For Ratings of Familiarity and Liking

Piece	Familiarity *M*	Familiarity *SD*	Liking *M*	Liking *SD*
Mozart	2.66	0.80	4.00	0.69
Webern	1.54	0.78	2.09	0.98
Wishart	1.57	0.74	2.86	1.00
Dean	1.49	0.56	2.51	1.04
Xenakis	1.43	0.70	2.09	1.04

Note. The familiarity scale ranged from 0: *I have never heard anything like this before* to 5: *I often listen to this piece of music* with *I have heard this piece* as the midpoint of the scale; the likeability scale ranged from 0: *Really dislike* to 5: *Really like* with *Neither like or dislike* as the midpoint of the scale.

Procedure

Participants first read an experiment information sheet, gave written informed consent, and received standardized instructions regarding the task. Specifically, participants were told they would be presented with five musical excerpts. Their task was to continuously rate their level of engagement throughout each piece of music. Drawing from the work of Schubert, Vincs, and Stevens (2013), we defined musical engagement as *compelled, drawn in, connected to what is happening in the music, interested in what will happen next*. Once participants received instructions and were clear with the definition, two practice trials were completed, followed by one block of the five experiment trials. After each trial, participants answered the familiarity and likability questions for that specific excerpt of music. After all experiment trials, participants completed the AIMS and the OMSI before debriefing. Overall, the experiment took ~30 min to complete.

Continuous Ratings of Perceived Arousal and Valence

In addition to ratings of continuous engagement measured from the sample of participants in the present study, perceived arousal and valence data from two previous experiments were used in our time series models. First, group mean perceived arousal and valence time series responses to the Webern, Wishart, Dean, and Xenakis excerpts (published in Bailes & Dean, 2012) were used in addition to perceived arousal and valence data in response to the Mozart excerpt (from a submitted paper). In this latter experiment, 22 first-year psychology participants ($M = 21.05$ years, $SD = 5.06$) continuously rated perceived arousal and valence using the identical two-dimensional emotion–space interface and procedure described in Bailes and Dean (2012) and in the present study.

Statistical Approach

For the purpose of time series analysis, stationarity of each series was assessed using the Augmented Dickey-Fuller Generalized Least-Squares test (Dickey & Fuller, 1979). Stationarity is achieved when a time series has constant mean and variance, which in turn means that autocorrelations (between events a certain number of time points apart) are also constant. It is required for secure interpretation by most time series analysis techniques. In each case here the time series could be "differenced" to stationarity. This creates a new series corresponding to differences between successive values of the original. A differenced linear time series without (measurement) error has a constant value, and differencing a time series that has a trend and substantial variability produces short runs of positive values, then negative values. As a result, such differenced series are commonly stationary even if the original series is not. A series resulting from differencing *seriesname* (e.g., arousal) is termed *dseriesname* (e.g., darousal). Autoregressive modeling with external predictors (ARX) was used for the analysis, as introduced in detail previously (Dean & Bailes, 2010). BIC was used as the basis of model selection: this penalizes strongly for the addition of predictor variables to a model (lowest BIC values are best) and hence provides a stringent test of whether engagement is a powerful predictor, even in conjunction with the acoustic predictors intensity and spectral flatness that represent aspects of loudness and timbre. When a pair of ARX models are compared and there is an absolute BIC difference ("delta BIC") of >4.6 between them, the evidence in favor of models with lower BIC is normally described as "strong" and correspond to a several-fold difference in probability. A delta BIC greater than 1.4 is termed "positive" in favor of the model with lower BIC (which shows increased probability), and smaller differences are ambiguous as to which model is preferred (Kass & Raftery, 1995).

We investigated whether engagement is influential in optimized univariate ARMAX models using acoustic predictors (as described previously for four pieces, but newly developed here for the Mozart piece). This means that a single perceptual variable (arousal or valence) was modeled with acoustic and engagement series as predictors, and optimized to investigate whether engagement contributed to the BIC-preferred model. Thus, a model including engagement with a delta BIC <−1.4 in relation to its acoustic predictor-only counterpart supports the tested hypothesis, and one with delta BIC <−4.6 very strongly supports it. For information, we show model comparisons in Table 4 wherever the delta BIC is negative, but they should be interpreted in relation to the two delta BIC limits just discussed. Only models that have white noise residuals (lacking autocorrelation) are considered here. The extent to which models fit the data is also reported.[2]

Results

Summary Features of the Five Pieces and the Responses to Them

Summary statistics for continuous ratings of engagement, arousal, and valence are displayed in Table 2, and corresponding statistics for the acoustic variables under consideration are displayed in Table 3. In each case, continuously measured engagement is strongly autoregressive, like the other perceptual variables (order up to 4, that is, up to four lags of the measure are predictive of the next). This is a likely feature both of the perception itself and of the motor response of moving the mouse in the computer-based interface, thus requiring the use of time series analysis rather than more common approaches that depend on data points being independent samples. Figure 1 illustrates the time course of grand average perceived engagement and arousal together, with the intensity profile for a select sample of pieces.

Table 2
Mean Ratings and Coefficients of Variation of Real-Time Perceived Engagement and Affect

Piece	Engagement		Arousal		Valence	
	M	*CV*	*M*	*CV*	*M*	*CV*
Webern	36.47	0.64	9.40	0.16	9.54	0.05
Wishart	62.83	0.33	11.89	0.19	−24.46	0.33
Dean	58.25	0.42	28.84	0.07	−18.84	0.11
Xenakis	34.81	0.62	14.83	0.04	−6.32	0.08
Mozart	73.69	0.20	23.99	0.42	32.69	0.16

Note. Ratings of perceived arousal and valence for Webern, Wishart, Dean, and Xenakis are taken from the experiment published in Bailes and Dean (2012); Ratings of perceived arousal and valence for Mozart are data from an unpublished experiment (see Method section for more detail). The Engagement scale ranged from 0 to 100; the Arousal and Valence scales ranged from −100 to +100.

Table 3
Means and Coefficients of Variation For Acoustic Intensity and Spectral Flatness

Piece	Intensity (dB)		Spectral flatness (Weiners)	
	M	*CV*	*M*	*CV*
Mozart	61.84	0.11	−9.88	0.13
Webern	49.27	0.20	−10.13	0.16
Wishart	52.45	0.25	−9.06	0.20
Dean	63.74	0.04	−3.73	0.12
Xenakis	60.53	0.04	−6.83	0.07

Time Series Analysis of Engagement as Granger Causal Predictor of Arousal and Valence

Hypotheses predicted either that engagement would be a strong predictor (H2), or a viable one (H1), and were investigated by first assessing whether the continuous parameter of engagement could enhance the best ARX models for darousal and dvalence separately, based solely on the optimal combination of acoustic parameters (dintensity and dspectral flatness). For all but Mozart, these are based on previous analyses reported in Bailes and Dean (2012). New ARX models for darousal and dvalence were developed for the Mozart piece. The predictors and results for the optimal models obtained here are shown in Table 4. Model results indicate whether dengagement was a predictor, in addition to the acoustic predictors of intensity and spectral flatness. This is a stringent modeling approach that considers whether the acoustic variables make the measure of engagement redundant, and also vice versa.

For darousal, engagement supported H1 for all but the Xenakis piece. It supported H2; that is, provided "strong" evidence for enhancement of the model based on acoustic parameters for the Wishart piece. The arousal profile in response to the Xenakis piece was the most poorly modeled (only about 17% of data squares are explained), and engagement is not beneficial as a predictor, as seen in Table 4 by its exclusion in significant models of darousal.

Turning to valence prediction, engagement again supported H1 for all but Xenakis and Dean, and supported H2 for the Wishart model. As noted in earlier work, valence is less well modeled than arousal in every case, and the difference is large for Webern, Dean, and Xenakis (the three least familiar pieces). The ARX analyses thus show almost complete support for H1 and partial support for H2.

Investigating the Role of Engagement With Mutual Influences of Perceived Arousal and Valence

It is worth considering how engagement interacts with the other perceptual parameters. Vector autoregression (VAR)

[2] See Bailes and Dean (2012) and Dean and Bailes (2010) for more detail on these standard methods of time-series analysis applied to music perception.

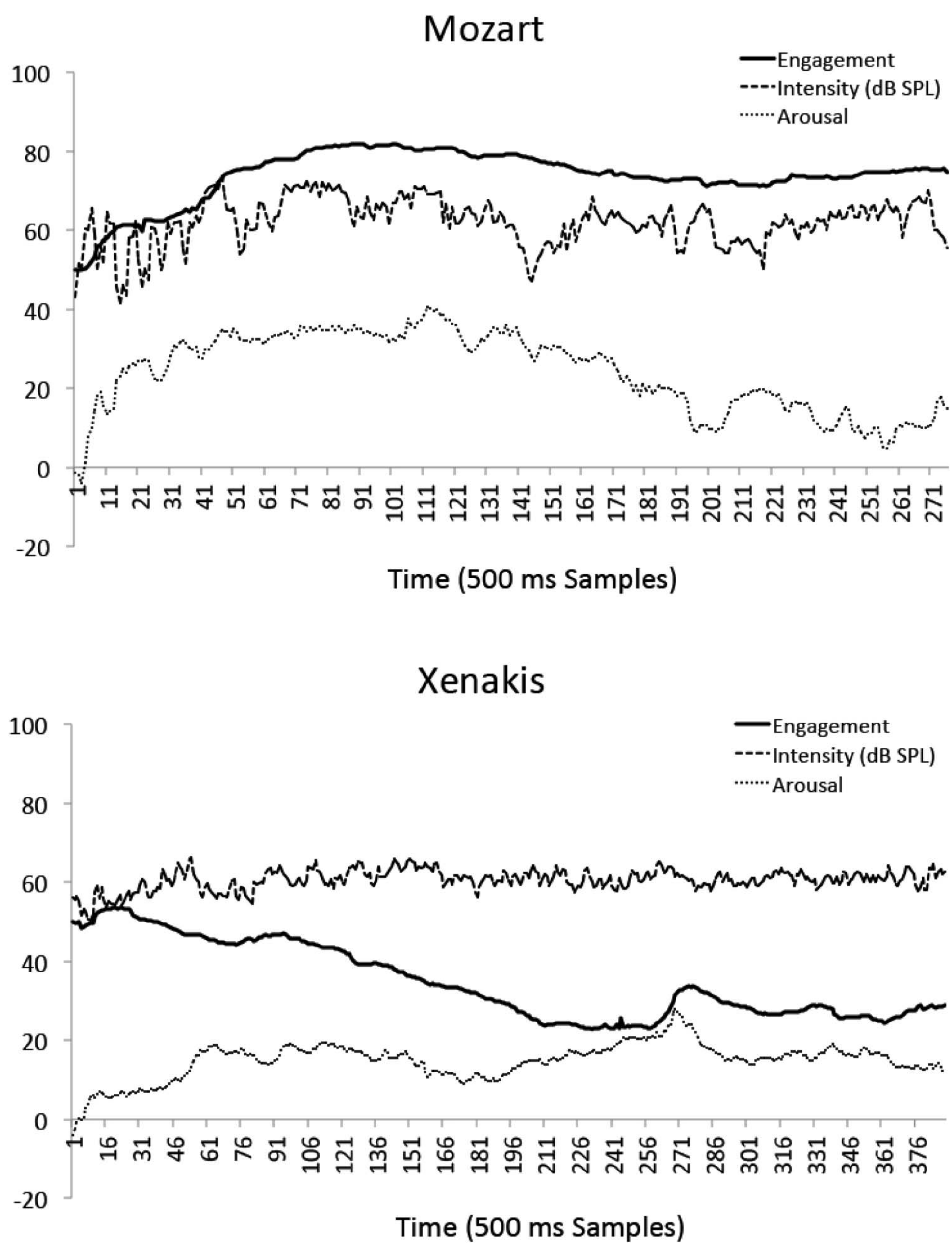

Figure 1. Mean engagement and arousal time series in response to Mozart and Xenakis, plotted with each excerpt's intensity profile (dB SPL). Engagement scale ranged from 0 (*Not Engaged*) to 100 (*Engaged*). Arousal scale ranged from −100 (*Very Passive*) to +100 (*Very Active*). Note that although the arousal scale ranged from −100 to +100, there was a clear tendency for participants to rate their perception on the "Active" half of the arousal dimension. The horizontal axis reports time and each data point was sampled at 500 ms (2 Hz).

analysis permits an assessment of the mutual impact of variables and their contribution to an overall model. Table 5 presents optimized VAR models based on the ARMAX models above, in addition to assessing whether all perceptual variables (perceived arousal, valence, and engagement) act together with mutual influence (as "endogenous" variables in statistical terminology). The acoustic variables, in contrast, are treated strictly as exogenous variables; that is, they do not influence each other, nor are they influenced by the perceptual variables (they are psychological "independent" variables). The models are optimized specifically from the perspective of prediction of darousal and dvalence. In other words, dengagement is not included unless it enhances the prediction within the VAR of at least one of these, and does not reduce the other comparably. For this comparison, the BIC cannot readily be used, in part because the removal of a predictor such as engagement signifies

Table 4
ARX Models of Perceived Affect, Including Engagement, and Exogenous (Acoustic) Variables

Piece	Modeled (stationary) variable	Model autoregressive and predictor inputs (lags)	BIC	Delta BIC	% of squared variable values predicted by the model
Mozart	darousal	dintensity(3,5), ar(1)	943.86		20.82
	darousal	dintensity(3,5) dengagement(1), ar(1)	943.65	−.21	22.50
	dvalence	dintensity(1,3), dspectral flatness(5)	880.26		16.61
	dvalence	dintensity(3), dspectral flatness(6), dengagement(0)	878.88	−1.38	17.04
Webern	darousal	dintensity(2,3,4), dspectral flatness(2), ar(1,5,9)	1478.68		43.13
	darousal	dintensity(2,3,4), dspectral flatness(2), dengagement(0,5), ar(1,5,9)	1477.74	−.94	44.23
	dvalence	dintensity(5), ar(1,2,4,5)	1091.82		6.73
	dvalence	dintensity(5), dengagement(3), ar(1,2,4,5)	1091.85	.03	7.69
Wishart	darousal	dintens(1–10), ar(1,3)	1206.59		47.19
	darousal	dintens(1–10), dengagement(0), ar(1)	1187.48	−19.11	55.37
	dvalence	dintensity(2,3), dspectral flatness(4), ar(1–3)	1255.86		45.19
	dvalence	dengagement(1,2,5), ar(1,2)	1213.36	−42.50	52.59
Dean	darousal	ar(1)	1022.92		23.43
	darousal	dengagement(4), ar(1)	1016.32	−6.60	23.47
	dvalence	dspectral flatness(4), ar(1,6,8)	999.35		14.42
	dvalence	Engagement not beneficial to model	—	—	—
Xenakis	darousal	dintensity(0,4), ar(1)	651.91		16.67
	darousal	Engagement not beneficial to model	—	—	—
	dvalence	dintensity (3,5), dspectral flatness (5), ar(2)	965.69		6.94
	dvalence	Engagement not beneficial to model	—	—	—

Note. The table presents predictor lags (in parentheses), and only models with white noise residuals (which were not autocorrelated). ar = autoregressive model components; BIC = Bayesian Information Criteria (lower values are better; note that values cannot be compared between pieces, only between models of a single piece). In each case, the best model of dvalence and darousal is shown without engagement. Presented underneath each of these results are the best models of darousal and dvalence when engagement is included as a predictor. The percentage data squared predicted is an estimate of the proportion of the data that is explained by the model, given that conventional R^2 values are not available (nor apt) for such time series. The models here, while based on the previous models of Bailes and Dean (2012), are different in their imposed restrictions from those used there because of our purpose to assess the possible influence of engagement. Here we did not permit lags > 10 (5 s) to enter the model, nor did we allow moving average error terms (which lack a clear rationale in this context) in addition to autoregressive ones. Conversely, each of the ARMAX models of Table 5 in Bailes and Dean (2012) was permitted to use only one acoustic predictor to fulfill the objectives of that study, whereas here we permit both to be used together, according to their effectiveness.

the removal of a dependent variable/endogenous variable, which is itself being modeled. Rather, the R^2 and related parameters are more informative, together with the significance values for the coefficients of the individual model components (not shown). For these VAR models, R^2 estimates are provided which, while not identical to the distributions of squares presented in Table 4, are still related and readily comparable. It can be seen in Table 5 that the relative values of R^2 are fairly similar to the relative values of predicted squares.

The VAR results indicate that engagement is retained as an endogenous variable for every piece, suggesting that its influence, though sometimes modest, does make a contribution. Engagement is quite well modeled itself, with R^2 ranging between 0.21–0.58 across all pieces. Valence remains poorly modeled in comparison with arousal, with the sole exception of the Wishart model (valence model $R^2 = 0.51$). The VAR analyses thus provide evidence that engagement supports H1 with respect to all five pieces.

After a VAR, impulse response functions (IRF) can be used to assesses a forecast-error variance decomposition; the impact of a unit change in each endogenous variable on a given output, lag by lag, that is independent of the impact point in the time series (Bailes & Dean, 2012). In the case of Wishart, Figure 2 shows that the IRF for a unit change of dengagement within the model produces a clear increment in darousal and dvalence. Moreover, it is dvalence whose confidence limits (shaded area in the right panel of Figure 2) separate from a zero impulse response at 2 s (four lags), confirming a "strong" impact and suggesting a special relevance of engagement to valence perception. This result provides further support for H2. For the other pieces, however, the IRF of dengagement did not reveal such significant responses.

Overall, a conservative conclusion from these VAR results is that engagement mediates some perceptual affective responses, but its impact in these particular pieces is generally modest (H1), with the possible exception of Wishart.

Discussion

The present study aimed to test the possible effects of engagement on real-time perception of arousal and valence by investigating the influence of listeners' continuous engagement with a particular musical piece. The results of the present study show a modest benefit in most cases when using a measure of musical engagement to predict perceived arousal or valence, even in conjunction with all other measured acoustic (intensity, spectral flatness) and perceptual (arousal, valence) parameters. Overall, the modest yet statistically significant role of listener engagement in models of arousal and valence suggest that it most likely mediates the relationship between acoustic parameters in music and listeners' affective responses.

Table 5
VARX Models With Perceptual Variables as Endogenous and Acoustic Variables As Exogenous

Piece: Model	Variable modeled	Parms	RMSE	R^2	χ^2	p-χ^2
Mozart: d.arous d.valen d.engage, exog(l(1,3,5).d.intens l(5).d.specf) lags(1,2)	D_arousal	10	1.24	.21	71.69	<.001
	D_valence	10	1.11	.09	28.19	<.001
	D_engage	10	0.37	.31	122.54	<.001
Webern: d.arous d.valen d.engage, exog(l(2,3,4).d.intens l(2).dspecf) lags(1,2)	D_arousal	10	1.64	.42	277.77	<.001
	D_valence	10	0.99	.07	29.40	<.001
	D_engage	10	0.35	.27	137.99	<.001
Wishart: d.arous d.valen d.engage, exog(l(1/ 7, 9, 10).d.intens) lags(1,2)	D_arousal	16	1.05	.54	447.20	<.001
	D_valence	16	1.05	.51	388.81	<.001
	D_engage	16	0.37	.58	529.29	<.001
Dean: d.arous d.valen d.engage, exog(l(4).d.specf) lags(1,2)	D_arousal	7	0.97	.26	124.88	<.001
	D_valence	7	0.96	.12	47.65	<.001
	D_engage	7	0.34	.22	99.73	<.001
Xenakis: d.arous d.valen d.engage, exog(l(0,5).d.intens l(5).d.specf) lags(1,2)	D_arousal	9	0.55	.17	76.28	<.001
	D_valence	9	0.83	.05	22.31	.01
	D_engage	9	0.34	.21	99.94	<.001

Note. Table presents VAR models (optimized for prediction of darousal and dvalence) of the five pieces, with all perceptual variables treated as endogenous, and all acoustic variables as exogenous. The model specifications (based on STATA code) are interpreted as follows. In each case there was no constant in the model, as variables were differenced. The endogenous variables (which may influence each other) are listed first, then after the comma, the exogenous variables and the level of lags/autoregression permitted. Models for the individual endogenous variables emerge, described by the number of parameters involved (Parms), their root mean square error (RMSE), R^2, and the χ^2 value testing whether the model could arise by chance and the corresponding p value.

One possible explanation for the modest explanatory power of musical engagement is that engagement is a multidimensional construct. For example, the level of engagement a listener experiences with a piece of music is probably linked to the level of enjoyment experienced and familiarity with each genre. Enjoyment has been shown to significantly predict discrete retrospective ratings of affective engagement given by audience members in a live concert setting (Thompson, 2006), although we found no equivalent evidence derived from our continuous data. However, the two pieces for which engagement had the least statistical impact—Dean and Xenakis—were indeed the two with the lowest familiarity. These two pieces also provided two of the three lowest ratings for liking. Real-time changes of listener interest in a piece of music could form another underlying component of musical engagement, and would not necessarily correlate with enjoyment. Take, for example, the relatively unfamiliar electroacoustic pieces presented in the present study. The unusual combination of inanimate sound sources with a variety of sound-warping effects is likely to elicit a high level of interest from the novelty of the sound, but enjoyment may remain stagnant or even decrease to a relatively low level because of the unusual sonic organization and often extreme, unexpected attributes. Both the Dean and Xenakis pieces are closely related to noise music, using quite static sustained textures in comparison with most acoustic music. The other distinct case in our data, the Webern piece, had both familiarity and liking ratings below those of Mozart and Wishart. This may be related to the repetitive rhythmic structure of the Webern piece (successive groups of notes or small chords sharing a single duration value with frequent gaps of that same duration) that provided another kind of stasis, even though played on a familiar Western instrument (the piano).

Furthermore, it may be a somewhat surprising result that individual differences such as one's propensity to be "musically absorbed" (as measured by the AIMS; Sandstrom & Russo, 2013) did not significantly correlate with ratings of engagement. Perhaps it is the case here that the chosen musical genres had an impact; that is, regardless of their level of musical absorption, participants may have found the choice of excerpts engaging because of the novelty of listening to a relatively unfamiliar excerpt and/or genre. A variety of pieces from more familiar and wide-ranging genres may shed greater light on this matter. It also remains to be seen whether effects of musical expertise may mediate these potential relationships; the sample in the present study was controlled to minimize the influence of differences in musical training, sophistication, and expertise.

An investigation of the relative importance of, and relationships between, variables that may contribute to listener engagement will not only create a better understanding of engagement as a motivational construct, but also provide stronger models of affect in the context of music. The results of listener engagement in the present study are promising, and with continued investigations of its underlying components, the relationship between musical engagement and perceived affect may be realized with greater clarity. It is clear here that in the context of complex auditory stimuli such as music, a variety of interrelations between variables exist. Engagement *and* arousal can mutually influence perceived valence in response to some pieces, and if engagement is indeed multidimensional, then its underlying components will most likely impact on arousal and valence in a way that is yet to be empirically determined. Future studies of arousal and valence may usefully address mutual relationships between such potentially interrelated variables when modeling real-time affective responses to music.

Overall, the present results suggest that continuous change in levels of engagement with a piece of music throughout a listening experience will bear further attention as a mediator of perceptual affective responses.

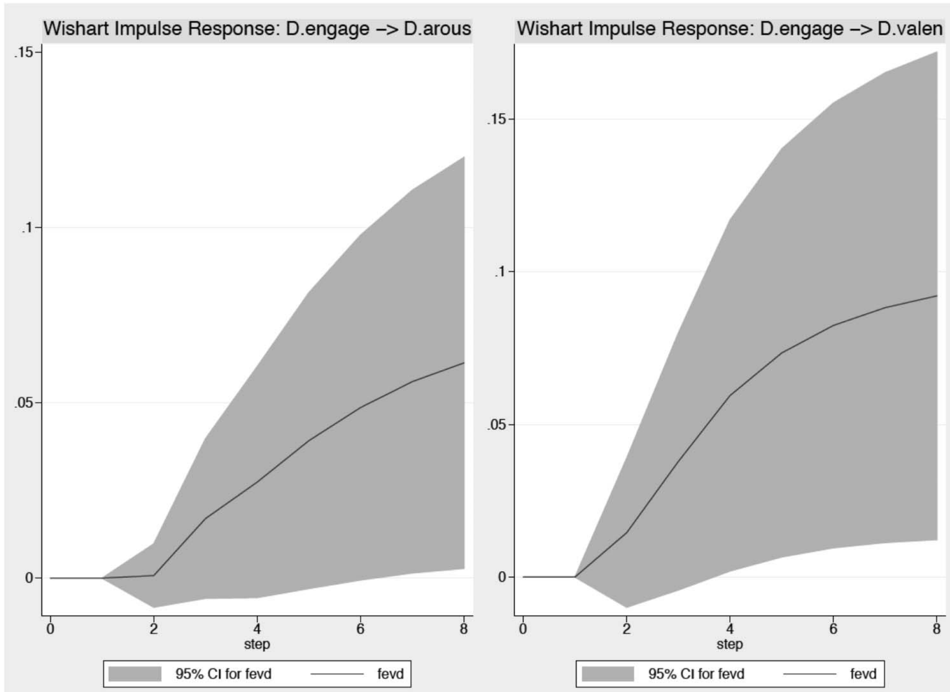

Figure 2. Forecast-error variance decomposition from a 4-lag (lag = 500 ms) VAR, representing the lag-by-lag impact of D.Engagement on D.Arousal (left panel) and D.Engagement on D.Valence (right panel) for the Wishart excerpt—the "impulse response functions." In particular, the confidence limits for D.Engagement -> D.Valence (shaded area in the right of panel) separate from a zero impulse response at 2 s (four lags), confirming a "strong" impact and suggesting a special relevance of engagement to valence perception.

References

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

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. doi:10.2307/40285811

Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research, 46,* 337–349. doi:10.1111/j.1468-5584.2004.00265.x

Bradley, M. M., & Lang, P. J. (2000a). Affective reactions to acoustic stimuli. *Psychophysiology, 37,* 204–215. doi:10.1111/1469-8986.3720204

Bradley, M. M., & Lang, P. J. (2000b). Measuring emotion: Behavior, feeling, and physiology. In R. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). New York, NY: Oxford University Press.

Brewster, A., Smith, H., & Dean, R. T. A. (2004). soundAffects. *Text, 8*(2). Retrieved from http://www.textjournal.com.au/oct04/smith2.htm

Broughton, M., Stevens, C. J., & Schubert, E. (2008). Continuous self-report of engagement to live solo marimba performance. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition* (pp. 366–371). Sapporo, Japan: Causal Productions.

Chin, T., & Rickard, N. S. (2012). The music USE (MUSE) questionnaire: An instrument to measure engagement in music. *Music Perception, 29,* 429–446. doi:10.1525/mp.2012.29.4.429

Dean, R. T., & Bailes, F. (2010). 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. (2011). Modelling perception of structure and affect in music: Spectral Centroid and Wishart's Red Bird. *Empirical Musicology Review, 6,* 131–137.

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

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time-series with a unit root. *Journal of the American Statistical Association, 74,* 427–431. doi:10.1080/01621459.1979.10482531

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music, 39,* 18–49. doi:10.1177/0305735610362821

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95,* 148–162. doi:10.1037/0022-0663.95.1.148

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367–400). Oxford: Oxford University Press.

Geringer, J. M., & Madsen, C. K. (2000–2001). A focus of attention model for meaningful listening. *Bulletin of the Council for Research in Music Education, 147,* 103–108.

Gingras, B., Marin, M. M., & Fitch, W. T. (2013). Beyond intensity: Spectral features effectively predict music-induced subjective arousal. *Quarterly Journal of Experimental Psychology,* 1–19. doi:10.1080/17470218.2013.863954

Gomez, P., & Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology, 53,* 91–103. doi:10.1016/j.ijpsycho.2004.02.002

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. doi:10.1525/mp.2007.24.3.297

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. doi:10.1525/mp.2006.23.4.319

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., & Sloboda, J. A. (2001). *Music and emotion: Theory and research*. Oxford: 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. doi:10.1017/S0140525X08005293

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90,* 773–795. doi:10.1080/01621459.1995.10476572

Korhonen, M. D., Clausi, D. A., & Jernigan, M. E. (2006). Modeling emotional content of music using system identification. *IEEE Transactions on Systems, Man, and Cybernetics*. Part B, Cybernetic: A Publication of the IEEE Systems, Man, and Cybernetics Society, *36,* 588–599. doi:10.1109/TSMCB.2005.862491

Leman, M., Vermeulen, V., De Voogdt, L., Moelants, D., & Lesaffre, M. (2005). Prediction of musical affect using a combination of acoustic structural cues. *Journal of New Music Research, 34,* 39–67. doi:10.1080/09298210500123978

Madsen, C. K., & Geringer, J. M. (2008). Reflections on Puccini's La Bohème: Investigating a model for listening. *Journal of Research in Music Education, 56,* 33–42. doi:10.1177/0022429408323072

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Ollen, J. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings*. (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.

Olsen, K. N., & Stevens, C. J. (2013). Psychophysiological response to acoustic intensity change in a musical chord. *Journal of Psychophysiology, 27,* 16–26. doi:10.1027/0269-8803/a000082

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion, 28,* 147–169. doi:10.1023/B:MOEM.0000032312.95499.6f

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. doi:10.1037/h0077714

Sandstrom, G. M., & Russo, F. A. (2013). Absorption in music: Development of a scale to identify individuals with strong emotional responses to music. *Psychology of Music, 41,* 216–228. doi:10.1177/0305735611422508

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. (2010). Continuous self-report methods. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 223–253). Oxford: Oxford Univeristy Press.

Schubert, E. (2013). Reliability issues regarding the beginning, middle and end of continuous emotion ratings to music. *Psychology of Music, 41,* 350–371. doi:10.1177/0305735611430079

Schubert, E., Vincs, K., & Stevens, C. J. (2013). Identifying regions of good agreement among responders in engagement with a piece of live dance. *Empirical Studies of the Arts, 31,* 1–20. doi:10.2190/EM.31.1.a

Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83,* 268–277. doi:10.1037/h0036681

Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). *Motivation and Emotion, 2,* 1–34. doi:10.1007/BF00992729

Thayer, R. E. (1986). Activation (arousal): The shift from a single to a multidimensional perspective. In J. Strelau, F. Farley, & A. Gale (Eds.), *The biological bases of personality and behavior* (Vol. 1, pp. 115–127). London: Hemisphere.

Thompson, S. (2006). Audience responses to a live orchestral concert. *Musicae Scientiae, 10,* 215–244. doi:10.1177/102986490601000204

Thompson, S. (2007). Determinants of listeners' enjoyment of a performance. *Psychology of Music, 35,* 20–36. doi:10.1177/0305735607068886

(*Appendix follows*)

Appendix

Accessible Recordings of the Musical Excerpts

1. Wolfgang Amadeus Mozart, *Piano Concerto 21, K. 467*: <https://www.youtube.com/watch?v=CoxL6uQjDUA>

2. Anton Webern (1937) *Piano Variations Op. 27*: <https://www.youtube.com/watch?v=ZEtqEzPakxA&list=RDZEtqEzPakxA#t=15>

3. Roger Dean (2003) *soundAFFECTS*: <http://www.textjournal.com.au/oct04/smith2.htm>

4. Trevor Wishart (1977) *Red Bird, a political prisoner's dream*: <http://www.ubu.com/sound/wishart.html>

5. Iannis Xenakis (1962) *Bohor*: <https://www.youtube.com/watch?v=-wo8LeaUK94> and <https://www.youtube.com/watch?v=-bq5GJ0pfTo>

Received February 13, 2014
Revision received April 17, 2014
Accepted April 22, 2014 ■