Simplifying the Implication-Realization Model of Melodic Expectancy

Article *in* Music Perception · April 1997

DOI: 10.2307/40285723

E. Glenn Schellenberg
University of Toronto

Simplifying the Implication-Realization Model of Melodic Expectancy
Author(s): E. Glenn Schellenberg
Reviewed work(s):
Source: *Music Perception: An Interdisciplinary Journal*, Vol. 14, No. 3 (Spring, 1997), pp. 295-318
Published by: University of California Press
Stable URL: http://www.jstor.org/stable/40285723
Accessed: 03/02/2012 21:19

Music Perception
Spring 1997, Vol. 14, No. 3, 295–318

Simplifying the Implication-Realization Model of Melodic Expectancy

E. GLENN SCHELLENBERG
University of Windsor

Results from previous investigations indicate that the implication-realization (I-R) model (Narmour, 1990) of expectancy in melody may be overspecified and more complex than necessary. Indeed, Schellenberg's (1996) revised model, with two fewer predictor variables, improved predictive accuracy compared with the original model. A reanalysis of data reported by Cuddy and Lunney (1995) provided similar results. When the principles of the I-R model were submitted to a principal-components analysis, a solution containing three orthogonal (uncorrelated) factors retained the accuracy of the model but was inferior to the revised model. A separate principal-components analysis of the predictors of the revised model yielded a two-factor solution that did not compromise the revised model's predictive power. Consequently, an even simpler model of melodic expectancy was derived. These results provide further evidence that redundancy in the I-R model can be eliminated without loss of predictive accuracy.

NARMOUR (1990, 1992) proposes a theory of melodic perception and cognition known as the *implication-realization* (I-R) model. Narmour's model is based on the idea that when listeners hear a melody they typically form expectancies about its continuation. These expectancies are presumed to result from a combination of innate and learned factors (see also Jones, 1990). The inclusion and specification of innate factors in the model is, arguably, one of its most important contributions to the psychology of music. The influence of these factors on melodic expectancies has been shown to generalize across listeners from different musical cultures tested in a wide variety of musical and nonmusical contexts (Cuddy & Lunney, 1995; Schellenberg, 1996), consistent with their proposed universality. Nonetheless, the I-R model appears to be overspecified and more complex than necessary (Schellenberg, 1996). The present report examines the degree to which the model can be simplified while retaining its predictive accuracy. Simplicity is one of the main criteria by which psychological models

Address correspondence to E. Glenn Schellenberg, Department of Psychology, University of Windsor, Windsor, Ontario, Canada N9B 3P4. (e-mail: schelle@uwindsor.ca)

are evaluated (e.g., Cutting, Bruno, Brady, & Moore, 1992). Indeed, the principle of parsimony (Occam's razor) is a scientific canon stating that a simpler explanation of any phenomenon is preferable to a more complex explanation. Simpler models, with fewer parameters, also "stand a better chance of being scientifically replicable" (Bentler & Mooijaart, 1989, p. 315). Hence, a simplified model of melodic expectancy is more likely to describe a universal phenomenon (i.e., one that could be replicated across all listeners in all musical contexts).

Learning is obviously involved in the expectancies that arise while listening to melodies. For example, a listener familiar with Western music (i.e., music from the common practice period, or from traditional or popular idioms) would generally expect a Western melody to stay "in key." The claim of innate components in melodic expectancy is more contentious. According to the I-R model, listeners, regardless of cultural background or exposure, form common expectancies that stem from innate psychological principles of perception and cognition. Any melodic interval that is not perceived as complete (or closed) is considered an *implicative interval*; such intervals imply that some tones are more likely than others to follow. (For a detailed account of factors causing closure, see Narmour, 1990.) The interval between the second tone of an implicative interval and the following tone is considered a *realized interval*. At its most basic (i.e., tone-to-tone) level, the I-R model describes how implicative intervals imply realized intervals and how some realized intervals are more implied than others. Narmour (1990, 1992) claims that these implications result from five perceptual predispositions (acting in combination with learned factors): *registral direction, intervallic difference, registral return, proximity,* and *closure*. A cursory description follows (for a more detailed description, see Schellenberg, 1996).

The first two predispositions (registral direction and intervallic difference) form the core of the I-R model; from these, the *basic melodic structures* are derived. Both principles depend on whether the implicative interval is small (five semitones or fewer) or large (seven semitones or more), because small intervals are considered to have different implications than large intervals. (The tritone [six semitones] is considered a threshold interval that is neither small nor large.) The principle of registral direction states that small intervals imply a continuation of pitch direction (e.g., a small upward implicative interval implies an upward realized interval), whereas large intervals imply a change of direction (e.g., a large upward implicative interval implies a downward realized interval or a unison). The principle of intervallic difference states that small implicative intervals imply similarly-sized realized intervals (defined as the same size ± two semitones if the realized interval changes registral direction, the same size ± three semitones otherwise), whereas large implicative intervals imply relatively smaller realized intervals.

The remaining three predispositions are not "principles" of the I-R model in the same way that the first two are. Nonetheless, all five are considered part of a "genetic code" (Narmour, 1989) for the perception of melodies. In keeping with earlier reports (Cuddy & Lunney, 1995; Schellenberg, 1996), all five will be considered principles. The principle of registral return describes a melodic archetype in which the second tone of the realized interval is proximate in pitch (by two semitones or fewer) to the first tone of the implicative interval, creating a symmetric (aba) or approximately symmetric (aba') pattern of pitches about a point in time. For example, the sequence C_4-G_4-C_4 is symmetric, whereas C_4-G_4-$C\sharp_4$ is approximately symmetric. Patterns become less archetypal (and less implied) as they deviate from exact mirror-image symmetry. The proximity principle states that small realized intervals (defined by Narmour as five semitones or fewer) are more implied than large intervals and that implications are stronger for relatively smaller realized intervals. For example, intervals of five semitones are more implied than intervals of six semitones, intervals of four semitones are more implied than five semitones, and so on. Finally, the principle of closure describes how listeners perceptually segment melodies as a function of pitch direction and interval size. Closure occurs when a melody changes direction (i.e., the implicative and realized intervals are in different directions) or when a relatively smaller realized interval follows a larger implicative interval; these two contributing factors are not mutually exclusive. All five principles apply to both upward and downward implicative intervals.

Schellenberg (1996) quantified each of the principles for 263 different combinations of implicative and realized intervals, represented by the grid in Figure 1. Only intervals one octave or smaller are considered in the grid; melodic intervals larger than octaves are relatively rare across musical cultures (Dowling & Harwood, 1986). Implicative intervals of 6 semitones (tritones) and 12 semitones (octaves) are also excluded because Narmour (1990, 1992) considers tritones to be a threshold interval and octaves a unique interval by virtue of octave equivalence. Each principle applies to both upward and downward implicative intervals, the grid making no distinction between pitch directions. Hence, each square in the grid (with one exception) represents two different combinations of implicative and realized intervals. For example, the square in the bottom left corner represents an upward implicative interval of 11 semitones followed by a downward realized interval of 12 semitones (e.g., C_4-B_4-B_3) or a downward implicative interval of 11 semitones followed by an upward realized interval of 12 semitones (C_4-$C\sharp_3$-$C\sharp_4$). Only the middle square in the top row, representing

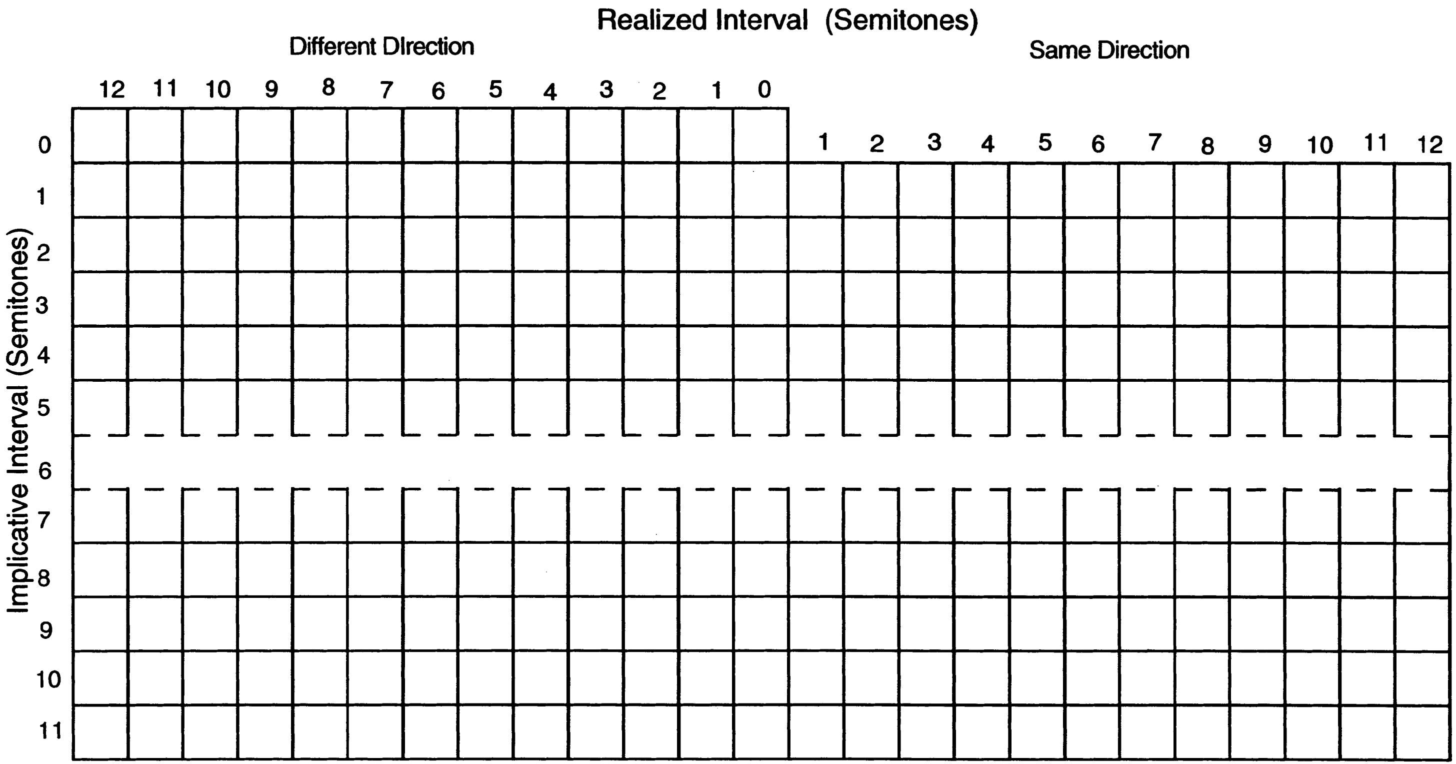

Fig. 1. A grid representing possible combinations of implicative and realized intervals. Because the principles of the implication-realization model apply to both upward and downward intervals, the grid does not distinguish between the two. The ordinate represents the size of the implicative interval, ranging from 0 to 11 semitones, and is subdivided into small (≤5 semitones) and large (≥7 semitones) intervals. The size of the realized interval is plotted on the abscissa, ranging from 12 semitones in a direction opposite to that of the implicative interval to 12 semitones in the same direction. Reprinted from *Cognition, 58*(1), E. G. Schellenberg, "Expectancy in melody: Tests of the implication-realization model," 75–125, 1996, with kind permission of Elsevier Science–NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

two consecutive unisons (an implicative interval of 0 semitones followed by a realized interval of 0 semitones), is a single combination.

Schellenberg (1996) quantified values for each of the five principles to predict listeners' responses in three experiments.[1] As shown in Figure 2, each pair of implicative and realized intervals was given a numerical value for each of the five principles. The first two principles (registral direction and intervallic difference) were coded as *dummy variables*. Instances where a realized interval conformed to the implication generated by the implicative interval were coded as 1; other instances (representing violations of the implications) were coded as 0. The other three principles (registral return, proximity, and closure) were coded such that instances of relatively greater conformity to the implications received relatively higher values. For example, the proximity principle was coded 6 for instances of maximum proximity (i.e., when the realized interval was a unison), 5 when the realized interval was one semitone, 4 when the realized interval was two semitones, and so on. Although these numerical values appear to be rather arbitrary, they represent accurately the principles of the I-R model (E. Narmour, personal communications, February 1990, April 1990, May 1991, June 1991).

In each experiment, Schellenberg's (1996) listeners heard fragments of melodies (piano timbre), each of which ended with an implicative interval, and rated (on a 7-point scale) how well individual test tones added to the end continued the fragments. The test tones were always within an octave of the last tone of each fragment. The fragments were from British folk songs (Experiment 1), Webern lieder (Experiment 2), and Chinese pentatonic folk songs (Experiment 3). Listeners in Experiments 1 and 2 varied in musical training, those in Experiment 3 varied in cultural background (American or Chinese). Standard multiple regression was used to model the outcome variable (listeners' ratings) as a function of five predictor variables (the principles of the I-R model). Potential experiential influences of tonality were minimized by including a tonality covariate in the analyses where appropriate (in Experiments 1 and 3). In general, the I-R principles successfully predicted response patterns, which did not differ as a function of musical style, formal musical training, or cultural background. In short, the results were consistent with Narmour's (1990, 1992) claim of innate influences on melodic expectancies. A summary of Schellenberg's (1996) results with the I-R model is provided in Table 1 (Model 1).

In an earlier study of melodic expectancy, Carlsen (1981) tested music students from Germany, Hungary, and the United States. The students heard two-tone stimulus intervals and were required to sing continuations of these

1. These experiments were initially conducted in collaboration with Carol L. Krumhansl. Krumhansl (1995) provided a summary of the data for a music audience. Schellenberg (1996) provided a complete account for a psychology audience.

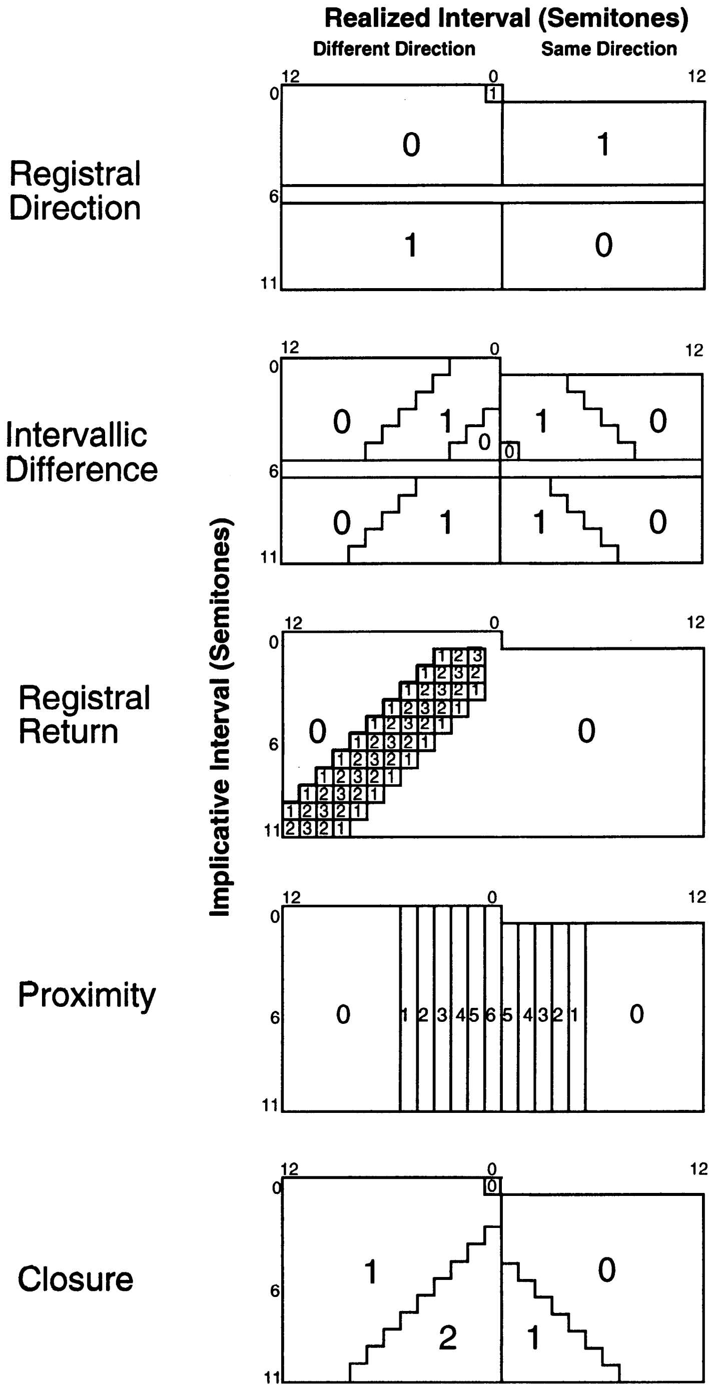

Fig. 2. Schellenberg's (1996) quantification of the five principles of the implication-realization model. Reprinted from *Cognition, 58*(1), E. G. Schellenberg, "Expectancy in melody: Tests of the implication-realization model," 75–125, 1996, with kind permission of Elsevier Science–NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

TABLE 1
Multiple Regression Results from Data of Schellenberg (1996, Appendixes A, B, and C)

	R	R^2	*F*	Degrees of Freedom Model	Degrees of Freedom Residual
Experiment 1 (*N* = 120)*					
Model 1: Implication-realization (I-R) model	.826	.683	40.49	6	113
Model 2: Revised model	.871	.759	90.65	4	115
Model 3: Principal-components model (I-R)	.818	.669	58.03	4	115
Model 4: Principal-components model (revised)	.871	.759	121.62	3	116
Model 5: Two-factor model	.871	.759	121.88	3	116
Experiment 2 (*N* = 200)					
Model 1: I-R model	.679	.461	33.18	5	194
Model 2: Revised model	.726	.527	72.73	3	196
Model 3: Principal-components model (I-R)	.669	.447	52.86	3	196
Model 4: Principal-components model (revised)	.724	.524	108.38	2	197
Model 5: Two-factor model	.721	.520	106.80	2	197
Experiment 3 (*N* = 132)*					
Model 1: I-R model	.830	.690	46.27	6	125
Model 2: Revised model	.869	.755	97.66	4	127
Model 3: Principal-components model (I-R)	.829	.688	70.00	4	127
Model 4: Principal-components model (revised)	.868	.754	130.62	3	128
Model 5: Two-factor model	.866	.750	127.89	3	128

Note. All *ps* < .0001. For each regression model, the number of predictor variables equals the degrees of freedom for the model.

*These models included a covariate to control for effects of tonality.

intervals. Although Carlsen was not attempting to test the I-R model, his stimuli can be considered implicative intervals. Intervals between the second tone of his stimuli and the first tone of responses can be considered realized intervals. Reanalyses of Carlsen's data revealed a good fit to the I-R model (Cuddy & Lunney, 1995; Schellenberg, 1996) and no differences as a function of cultural background (Schellenberg, 1996). A reanalysis of Unyk and Carlsen's (1987) replication of the study by Carlsen (1981) with American music students also revealed that response patterns were consistent with predictions of the I-R model (Schellenberg, 1996).

Nonetheless, strong correlations among some of the quantified principles of the I-R model raise the possibility of redundant predictors. Pairwise correlations between predictors are provided in Table 2, calculated for the 263 combinations of implicative and realized intervals shown in Figure 1. The correlations reveal that intervallic difference, proximity, and closure form a highly collinear set of predictors. Thus, inclusion of all three principles in the model may be unnecessary.

Schellenberg (1996)'s revised model (Table 1, Model 2), which contained two fewer principles than the original I-R model (Table 1, Model 1), in fact predicted his listeners' response patterns with greater accuracy. Quantified

TABLE 2
Pairwise Correlations between Principles of the Implication-Realization Model and the Revised Model ($N = 263$)

IMPLICATION-REALIZATION MODEL	Intervallic Difference	Registral Return	Proximity	Closure
Registral direction	.144*	.032	.014	.010
Intervallic difference		.035	.628**	.320**
Registral return			–.028	.125*
Proximity				.380**
REVISED MODEL	**Registral Return–Revised**		**Proximity Revised**	
Registral direction–revised	.336**		–.047	
Registral return–revised			.004	

*$p < .05$.
**$p < .0001$.

values for the three revised principles are provided in Figure 3. Because of intercorrelations among the original principles of intervallic difference, proximity, and closure, Schellenberg omitted intervallic difference and closure and revised proximity, forming a new predictor variable that he designated *proximity-revised*. Proximity-revised simply represents the size of the realized interval (in semitones); intervals are less implied as they become larger. The remaining two principles were also modified. Registral direction was revised so that it applies only to large intervals, with large implicative intervals implying a realized interval in a different direction. Finally, Schellenberg revised registral return as an all-or-none (dummy) variable, eliminating the distinction between symmetric and approximately symmetric pitch patterns. Although Schellenberg derived his revised principles on the basis of the data from the first of his three experiments, the revised model proved superior to the original at predicting listeners' responses across all three experiments.

Reanalysis of Data from Cuddy and Lunney (1995)

Cuddy and Lunney (1995) used a task similar to that of Schellenberg (1996), the principal difference being that their listeners were required to rate test tones as continuations of two-tone implicative intervals rather than fragments of real melodies. Their intervals were major seconds, minor thirds, major sixths, and minor sevenths, each of which was presented in ascending and descending forms (a total of eight different intervals) with a piano timbre. For each interval, listeners rated all 25 chromatic tones within

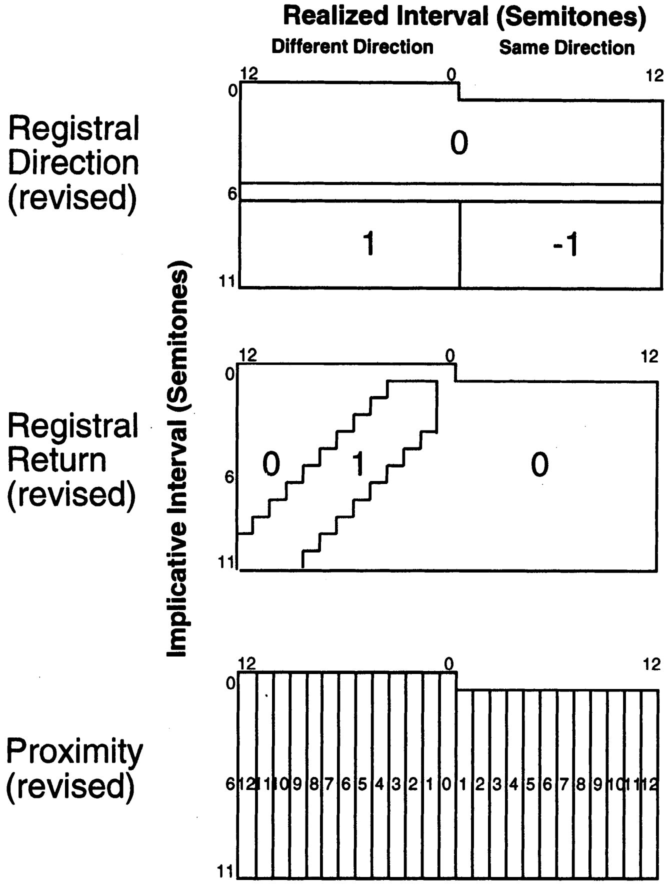

Fig. 3. Schellenberg's (1996) quantification of the three principles of his revised model. Reprinted from *Cognition, 58*(1), E. G. Schellenberg, "Expectancy in melody: Tests of the implication-realization model," 75–125, 1996, with kind permission of Elsevier Science–NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.

an octave from the second tone of the implicative interval (total of 200 ratings). The first tone of each interval was a dotted quarter note, the second tone was an eighth note, and the test tone was a quarter note, such that durations stood in a ratio of 3:1:2. Again, standard multiple regression analyses revealed that the I-R model successfully predicted response patterns well above chance levels even though the stimulus contexts were musically impoverished.

On the basis of their findings, Cuddy and Lunney (1995) revised registral direction so that it applied only to large intervals, although their method of requantification was different from that of Schellenberg (1996). They also used the revised version of registral return as coded in Figure 3. Their most complete explanation of response patterns was provided by a multiple regression model that contained a revised version of the I-R model (closure was excluded, registral direction and registral return were revised) and three covariates (one controlling for variation in ratings due to pitch height, the other two controlling for influences of tonality).

Because Cuddy and Lunney (1995) did not use the revised model as coded here, an initial analysis provided a direct comparison of the revised and original I-R models. The results are summarized in Table 3 (Models 1 and 2). As with Schellenberg's (1996) data, the revised model did not result in a loss of predictive accuracy compared with the original I-R model. Rather, 72.5% of the variance was explained by the revised model (plus the three covariates) compared with 64.0% for the I-R model (plus covariates).

To summarize, the I-R model can successfully predict ratings in melodic expectancy tasks well above chance levels regardless of differences in musical styles and groups of listeners. Nonetheless, considered in combination with Schellenberg's (1996) findings, the reanalysis of Cuddy and Lunney's (1995) data confirms that the I-R model can be substantially simplified without loss of predictive power.

Methods and Objectives

The objective of the present report was to derive the simplest possible model of expectancy in melody while retaining the predictive power of more complex models. Specifically, principal-components analysis was used to measure the number of nonredundant (i.e., orthogonal, or uncorrelated) dimensions embodied in the I-R model. A "principal-components" model consisting solely of these unique dimensions was evaluated for relative effi-

TABLE 3
Multiple Regression Results from Data of Cuddy and Lunney (1995, Appendix)

				Degrees of Freedom	
	R	R^2	F	Model	Residual
Model 1: Implication-realization (I-R) model	.800	.640	42.38	8	191
Model 2: Revised model	.851	.725	84.67	6	193
Model 3: Principal-components model (I-R)	.794	.630	54.74	6	193
Model 4: Principal-components model (revised)	.849	.721	100.04	5	194
Model 5: Two-factor model	.851	.724	101.53	5	194

All ps < .0001 (N = 200). For each regression model, the number of predictor variables equals the degrees of freedom for the model. Each model included three covariates. Two of these held constant differences in responding due to the tonal implications of the stimulus patterns; the third controlled for influences of the pitch height of the test tones.

cacy in predicting the data of Cuddy and Lunney (1995) and those of Schellenberg (1996); the principal-components model was also compared with the original I-R model and Schellenberg's revised model. Finally, a separate principal-components analysis was conducted on the principles of the revised model; from this analysis, a further simplified model of melodic expectancy was derived.

PRINCIPAL-COMPONENT ANALYSIS

A brief description of principal components analysis follows for the benefit of readers unfamiliar with multivariate statistical techniques. (For a detailed account, see Tabachnick and Fidell [1996, chap. 13]; see Diekhoff [1992, chap. 16] for a concise overview. The current summary is derived from both sources.) Principal-components analysis is one of two statistical methods generally referred to as *factor analysis* (the other being the *common factor model*). In factor analysis, the goal is to reduce a set of variables (e.g., the principles of the I-R model) to a smaller set of orthogonal (uncorrelated) variables. These uncorrelated variables are called *factors*, each of which is a different linear combination of the original variables, derived such that each factor is uncorrelated ($r = 0$) with every other factor. Hence, factors are *latent variables,* not actually present in the set of original variables. Rather, they represent the underlying "structure" of the original set (i.e., the number of orthogonal dimensions contained in the set).

The first factor in a principal-components solution is the linear combination of original variables that accounts for the largest proportion of the variance in the entire set; this proportion is measured with an *eigenvalue.* A factor's eigenvalue is the sum of the squared correlations between the factor and each original variable. Thus, the first factor is the linear combination of original variables that has the largest possible eigenvalue. The second factor is a different linear combination of the original variables that is orthogonal to the first factor and explains the next largest proportion of variance in the original variables (having the second largest eigenvalue). The third factor is the linear combination orthogonal to the first and second factors that explains the next largest proportion of variance (having the third largest eigenvalue), and so on.

To explain *all* of the variance in a set of variables, one needs as many factors as there are variables unless some of the variables are perfectly correlated. Typically, however, the first few factors explain a large proportion of variance in the original set and are considered to be an adequate account of the set's structure (i.e., an adequate solution). There are no hard and fast rules by which to determine the "correct" number of factors to include in a solution. Nonetheless, a factor with an eigenvalue of less than 1 explains less variance than the variance found in a single original variable. Because the goal is to reduce the number of variables, it is common to constrain a solution such that only factors with eigenvalues greater than or equal to 1 are included.

Factors retained in a solution are rotated to increase their interpretability. Varimax rotation is the most common method. Rotation does not affect the orthogonality of the factors, but makes the solution more interpretable by increasing high pairwise correlations between factors and original variables and by decreasing low correlations. After rotation, it is easier to identify which variables are most strongly associated with each factor.

Principal-Components Analysis of I-R Model

The five principles of the I-R model were submitted to a principal-components analysis with varimax rotation to determine (1) whether a simplified version of the I-R model (i.e., no redundancy among predictors) retained its predictive accuracy and (2) whether the solution was similar to the revised model proposed by Schellenberg (1996). The data for this analysis consisted solely of the quantified values of the predictor variables provided in Figure 2 (i.e., the analysis did *not* involve the outcome variable—listeners' ratings). The analysis included all 263 combinations of implicative and realized intervals represented by the grid in Figure 1. Factors with eigenvalues less than 1 were excluded from the solution, which contained three factors. The results of the analysis are summarized in Table 4. The first factor, which accounted for the largest proportion of the variance (eigenvalue of 1.90), explained 37.9% of the variance in the original five principles. The second and third factors (eigenvalues of 1.04 and 1.02, respectively) accounted for 20.9% and 20.5% of the variance, respectively. Thus, 79.3% of the variance in the five principles of the I-R model can be explained by three orthogonal factors. The first factor was highly correlated with the principles of proximity, intervallic difference, and closure. The second factor was highly correlated with registral return and moderately correlated with closure. The third factor was highly correlated with registral direction and exhibited small but significant correlations with intervallic difference and closure.

The next set of analyses examined how well the three orthogonal factors from the principal-components analysis predicted melodic expectancy data from previous investigations (Cuddy & Lunney, 1995; Schellenberg, 1996). Hence, listeners' ratings were modeled as a function of three latent variables (i.e., the factors identified in the principal-components analysis) rather than the predictor variables specified by the I-R model. These analyses determined whether the proportion of variance in the I-R model left unex-

TABLE 4
Results from Principal-Components Analysis of the Implication-Realization Model

Principle	Coefficient	Standardized Coefficient	Correlation with Factor
Factor 1			
Registral direction	–.068	–.034	.037
Intervallic difference	.876	.436	.835***
Registral return	–.055	–.043	.001
Proximity	.242	.472	.876***
Closure	.495	.342	.655***
Constant	–1.079		
Factor 2			
Registral direction	.024	.012	.032
Intervallic difference	–.157	–.078	–.041
Registral return	1.200	.925	.962***
Proximity	–.072	–.140	–.107
Closure	.412	.284	.322***
Constant	–.554		
Factor 3			
Registral direction	1.922	.962	.983***
Intervallic difference	.308	.153	.201**
Registral return	.032	.024	.042
Proximity	–.031	–.061	–.016
Closure	–.244	–.168	–.130*
Constant	–.828		

*$p < .05$
**$p < .005$
***$p < .0001$

plained by the principal components solution (i.e., about one-fifth of the total variance) was necessary to explain listeners' response patterns. As shown in Tables 1 and 3, the principal-components (I-R) model (Model 3) was comparable in goodness-of-fit to the I-R model (Model 1) across data sets. In each case, substituting the three orthogonal factors from the principal-components (I-R) model for the five predictors of the I-R model resulted in a loss of predictive accuracy of less than 2%. Thus, although the three orthogonal factors leave approximately one fifth of the variance in the I-R model unexplained, this unexplained variance is relatively unimportant in explaining ratings of melodic expectancy.

Similarities between the principal-components (I-R) model and the revised model were examined by measuring pairwise correlations between the orthogonal factors and the revised principles. Each factor from the principal-components (I-R) model was most highly correlated with a different revised principle: the first factor with the revised proximity principle ($r = -.86$, $p < .0001$), the second factor with the revised registral return principle ($r = .89$, $p < .0001$), and the third factor with the revised registral direction principle ($r = .59$, $p < .0001$). These correlations reveal that the principles of the revised model, which was derived from listeners' expectancy ratings, are associated with those of the mathematically derived principal components (I-R) model. Nonetheless, the revised registral direction principle was also weakly correlated with the first and second factors ($r = .23$, $p < .0005$, and $r = .44$, $p < .0001$, respectively). Moreover, the revised model (Model 2) consistently outperformed the principal-components (I-R) model (Model 3) across the analyses reported in Tables 1 and 3, explaining between 5.3% and 9.5% more variance in each case. Hence, the revised model is not simply a reflection of the latent structure of the original I-R model. Rather, the changes incorporated into the revised model actually improve its predictive power relative to the original I-R model. Subsequent analyses focused on potential simplification and improvement of the revised model.

Principal-Components Analysis of the Revised Model

Although Schellenberg (1996) found that his revised model of melodic expectancy explained response patterns better than the I-R model, the revised model also contains some redundancy. As shown in Table 2, the revised principles of registral direction and registral return are correlated, implying that inclusion of both principles may be unnecessary. Accordingly, a second principal-components analysis was used to derive an even simpler model (the principal-components [revised] model) with no redundancy and no loss of predictive power.

As with the I-R model, the three principles of the revised model were submitted to a principal-components analysis with varimax rotation (same 263 combinations of implicative and realized intervals, same eigenvalue criterion); the results are summarized in Table 5. A solution containing two orthogonal factors was obtained. The first factor explained 44.5% of the variance in the three principles of the revised model (eigenvalue of 1.34), the second factor explained 33.5% (eigenvalue of 1.00). Hence, the two-factor solution explained 78.0% of the total variance in the revised model. High correlations with the revised principles of registral direction and registral return revealed that the first factor was essentially a linear combination of these two principles. By contrast, the second factor was almost perfectly correlated with the revised proximity principle.

The two factors uncovered in the principal-components solution were used to predict the data collected by Cuddy and Lunney (1995) and by Schellenberg (1996). As indicated in Tables 1 and 3, substitution of the two orthogonal factors from the principal-components (revised) model (Model 4) for the three predictors of the revised model (Model 3) resulted in a loss

TABLE 5
Results from Principal-Components Analysis of the Revised Model

Principle	Coefficient	Standardized Coefficient	Correlation with Factor
Factor 1			
Registral direction–revised	.880	.608	.814*
Registral return–revised	1.618	.616	.820*
Proximity–revised	.004	.013	–.013
Constant	–.323		
Factor 2			
Registral direction–revised	–.099	–.068	–.087
Registral return–revised	.219	.083	.065
Proximity–revised	.274	.993	.996*
Constant	–1.740		

*$p < .0001$.

of predictive accuracy of 0.3% or less. Thus, the proportion of variance in the revised model that is unexplained by the principal-components analysis (about one fifth of the total variance) is not needed to explain the available data. Moreover, these results imply that a model limited to two factors or principles might provide the best account of processes underlying the formation of tone-to-tone expectancies when listening to melodies.

A Simplified Model of Expectancy in Melody

The two-factor model of melodic expectancy is the outcome of the results reported in the preceding sections. Quantified values of the model's two principles are provided in Figure 4. The principles of the two-factor model were obtained by modifying the factors from the principal-components (revised) model such that they would be both theoretically relevant and easy to use. Thus, the two-factor model is essentially a validation of the principal-components (revised) model.

The first principle of the two-factor model is *pitch proximity.* This principle states that when listeners hear an implicative interval in a melody, they expect the next tone to be proximate in pitch to the second tone of the implicative interval (i.e., they expect a small realized interval). The pitch-proximity principle is identical to Schellenberg's (1996) proximity-revised principle (Figure 3), which was found to be almost identical ($r \cong 1$) to the second factor in the principal-components analysis of his revised model. The principle, which measures the size of the realized interval in semitones (unison = 0, minor second = 1, major second = 2, etc.), is simple to quantify. Because larger intervals are assigned higher values, *negative* associa-

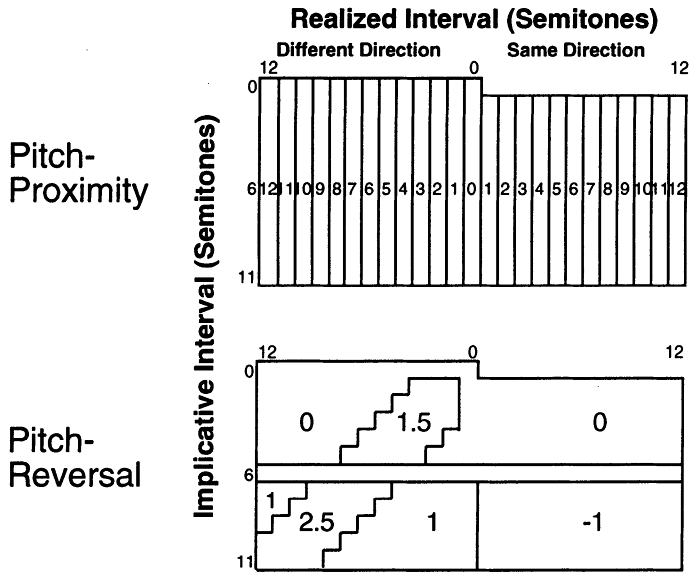

Fig. 4. Quantification of the principles of the two-factor model.

tions between the principle and expectancy data are expected. Although the principle is specified here using semitones, any logarithmic pitch scale could be substituted. Thus, the pitch-proximity principle can be used with scales whose intervals are not multiples of semitones. With the Thai scale, for example, where consecutive tones of the seven-tone scale are equidistant or approximately equidistant (Ellingson, 1992; Morton, 1980), one would code a unison as 0, an interval of one scale step as 1, an interval of two scale steps as 2, and so on. Moreover, the pitch-proximity principle does not suffer from the arbitrary assumption of Narmour's (1990, 1992) original concept of proximity, in which all realized intervals larger than five semitones are considered to be equally nonproximate (see Figure 2).

Expectancies are typically considered to represent learned schemas, whereas perceptual grouping based on proximity is considered to reflect a "primitive" (unlearned) process (Bregman, 1990). Nonetheless, because proximity tends to predict grouping in vision (Koffka, 1935; Kohler, 1947) and audition (Bregman, 1990), the pitch-proximity principle could stem from a hard-wired perceptual predisposition that also influences expectancies when listening to melodies, as Narmour claims. For example, listeners perceive tones that are proximate in pitch to be similar (Kallman, 1982). Moreover, research on "auditory scene analysis" (Bregman, 1990) indicates that proximity is a critical factor in the perception of auditory patterns that appears to be independent of learning. Tones proximate in pitch tend to be heard as originating from the same source or object; nonproximate tones are typically interpreted as originating from different sources. Thus, a melody could be relatively incoherent (i.e., difficult to process and represent; see Bharucha & Pryor, 1986; Schellenberg & Trehub, 1996) if its component tones are heard as emanating from multiple sources. Indeed, Deutsch (1978) reported more efficient processing of sequences with small intervals than of sequences with larger intervals.

The pitch-proximity principle can also be interpreted as a principle of conjunct motion (i.e., motion by scale step; see Aldwell & Schachter, 1989). The present results demonstrate that listeners typically expect melodies to continue in scale steps (the smallest possible intervals in a given mode) rather than in leaps (larger intervals). When a melody departs from conjunct motion, smaller leaps in pitch are less unexpected than larger leaps. It remains unclear, however, whether listeners' expectancies for proximate tones are innate rather than learned. Because small intervals are common in melodies (Dowling & Harwood, 1986), listeners could learn to expect typically sized intervals. Nonetheless, the cross-cultural predominance of small intervals in music (Dowling & Harwood, 1986) is consistent with the idea that the pitch-proximity principle describes a musical universal. Further, proximity is a robust predictor of responses from listeners from different musical cultures (Carlsen, 1981; Schellenberg, 1996), tested with various musical styles (Schellenberg, 1996) or in relatively nonmusical contexts (Carlsen, 1981; Cuddy & Lunney, 1995).

The second principle of the two-factor model, *pitch reversal,* is based on the first factor obtained in the principal-components analysis of the revised model (Table 5). This factor was essentially a linear combination of the revised principles of registral direction and registral return (i.e., Factor 1 = 0.880 [registral direction-revised] + 1.618 [registral return-revised] + 0.004 [proximity-revised] – 0.323, see Table 5); both revised principles were highly correlated with this factor. For ease of use, the coefficients for the revised principles of registral direction and registral return were rounded (from 0.880 to 1 and from 1.618 to 1.5, respectively), and the relatively small constant and proximity-revised coefficient were ignored, giving the following formula:

pitch-reversal = registral direction-revised + (1.5) registral return-revised.

When quantified in this manner, the pitch-reversal principle remains almost perfectly correlated with the first factor from the principal components analysis of the revised model ($r = .997$, $p < .0001$). Moreover, it is uncorrelated with the pitch-proximity principle ($r = .029$).

As illustrated in Figure 4, pitch reversal can have one of five values: –1, 0, 1, 1.5, and 2.5. The greater range of values for large implicative intervals (–1, 1, and 2.5) than for small implicative intervals (0 and 1.5) indicates that the principle is a relatively stronger determinant of expectancies for large intervals. Although the pitch-reversal principle is simply an additive combination of two principles from the revised model (registral direction-revised and registral return-revised), it is justified for two reasons: (1) it collapses two overlapping (correlated) principles—both of which specify a reversal of pitch direction—into a single principle, and (2) it results in a very parsimonious model of melodic expectancy with *no* redundancy among predictors.

The pitch-reversal principle extends the pitch-proximity principle to relations between noncontiguous tones. In addition to expecting the next tone in a melody to be proximate in pitch to the tone heard most recently, the principle claims that listeners often expect the next tone to be proximate in pitch to the tone that preceded the most recently heard tone. In other words, listeners often expect the second tone of a realized interval to be proximate to the first tone of the implicative interval. Hence, the pitch-reversal principle describes expectancies for proximate pitch relations slightly more global than those described by the pitch-proximity principle.

The pitch-reversal principle also describes expectancies that arise when a melody violates the pitch-proximity principle (i.e., when a large implicative interval is heard). Once the coherence of a melody has been "threatened" by *disjunct motion* (a melodic leap), the principle asserts that listeners expect a reversal of pitch direction. If this expectancy is considered jointly with the expectancy for small intervals (as described by the pitch-proximity principle), the overall expectancy is for the resulting gap in pitch to be filled (i.e., listeners expect a change of direction *and* a relatively small interval). Filling the gap helps to restore the integrity and coherence of the melodic line. An expectancy for "gap-fill" melodic patterns was articulated explicitly by Meyer (1973) and is consistent with rules of music theory and voice leading (e.g., Aldwell & Schachter, 1989). Moreover, previous investigations (Rosner & Meyer, 1982, 1986; Schmuckler, 1989) have provided evidence supporting the psychological validity of gap-fill expectancies.

Human vocal limitations may be another factor contributing to both principles of the two-factor model. Large intervals are more difficult to sing accurately than are small intervals (Bregman, 1990), so the pitch-proximity principle could stem, in part, from an expectancy for intervals that are easily sung. Moreover, if a tone following a large interval in a melody does not change direction, it is more likely to exceed a singer's vocal range than a tone that changes direction. Hence, the expectancy for pitch reversals could reflect knowledge of vocal range limitations.

Analyses involving predictions of data collected by Cuddy and Lunney (1995) and Schellenberg (1996) by means of the two-factor model are summarized in Tables 1 and 3 (Model 5). Because the model was derived from the principal-components analysis of the revised model, its fit to the various data sets was virtually identical to that of the principal-components (revised) model (Tables 1 and 3, Model 4), which, in turn, was almost identical in predictive power to the revised model (Tables 1 and 3, Model 2). Collapsing two principles (registral direction-revised and registral return-revised) into one (pitch-reversal) meant that the coefficients for the principles could not vary freely for each data set, making it impossible for the two-factor model to exceed the revised model in predictive power. Nonetheless, the loss of a degree of freedom had little consequence for predictive accuracy across data sets. This finding provides empirical validation of the pitch-reversal principle in particular and the two-factor model in general.

Although the two-factor model is substantially different from the original I-R model, it can still be considered a revision of the original. Both factors are based on ideas articulated by Narmour (1990, 1992). Moreover, some aspects of the two-factor model reflect central tenets of the I-R model. For example, the pitch-reversal principle retains Narmour's distinction between small and large intervals, his proposal that large intervals imply a reversal of pitch direction, and his suggestion that pitch patterns symmetric about a point in time are melodic archetypes. Nonetheless, the two-factor model can be considered superior to the original on the basis of its greater simplicity in accounting for melodic expectancies. Two of the principles from the original I-R model (i.e., intervallic difference and closure) appear to have little psychological validity, and none of the remaining principles as originally formulated is retained in the two-factor model. Indeed, the present report makes it clear that the original I-R model provides neither the most complete nor the most elegant description of the available data.

It is important to note, however, that the stimulus contexts examined in the present study were severely constrained. Specifically, the contexts and analyses focused on pitch relations at very local (tone-to-tone) levels, with rhythmic factors held constant. Rhythm plays a major role in determining expectancies (Jones, 1981, 1982, 1987, 1990; Jones & Boltz, 1989), and several studies have reported that rhythmic factors interact with those based on pitch (e.g., Boltz, 1991, 1993; Jones, Boltz, & Kidd, 1982; Jones, Summerell, & Marshburn, 1987). Because influences of pitch proximity were found to extend to noncontiguous tones (i.e., tones separated by one tone), more global properties of melodies are also likely to influence listeners' expectancies (Schmuckler, 1990).

In sum, the two-factor model proposes that tone-to-tone expectancies are determined primarily by proximity; upcoming tones in a melody are expected to be proximate to tones heard previously. When listeners hear successive tones that are nonproximate (relatively distant in pitch), they expect the next tone to fill in the gap. Despite the success of this very simple nonredundant model across data sets, it is conceivable that a more complex model with built-in redundancies might be required to explain melodic expectancies in other contexts. For example, Narmour (1992) claims that "the redundancy built into our melodic cognitive systems" (p. 239) helps to "eliminate perceptual processing 'mistakes'" (p. 244). Hence, the present focus on parsimony may have resulted in a model that is applicable only to contexts similar to the ones examined here, where "processing mistakes" may be relatively unlikely. Moreover, a melody's rhythmic and global pitch characteristics are likely to interact with the local pitch principles of the two-factor model. Nonetheless, the two-factor model provides a very simple and elegant explanation of the results from the studies examined in the present report. Future research could examine the predictive power of the two-factor model in other contexts. Interested researchers should bear in mind, however, that for stimulus contexts with tonal implications, covariates for influences of tonality may be required. Developmental and cross-cultural approaches would be particularly useful in assessing whether the two-factor model is universally applicable and whether learning and exposure to music are necessary for the principles to become operative.[2]

References

Aldwell, E., & Schachter, C. (1989). *Harmony and voice leading* (2nd ed.). San Diego: Harcourt Brace Jovanovich.

Bentler, P. M., & Mooijaart, A. (1989). Choice of structural model via parsimony: A rationale based on precision. *Psychological Bulletin, 106*, 315–317.

Bharucha, J. J., & Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception & Psychophysics, 40*, 137–141.

Boltz, M. (1991). Some structural determinants of melody recall. *Memory & Cognition, 19*, 239–251.

Boltz, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics, 53*, 585–600.

Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.

Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology, 1*, 12–29.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception & Psychophysics, 57*, 451–462.

Cutting, J. E., Bruno, N., Brady, N. P., & Moore, C. (1992). Selectivity, scope, and simplicity of models: A lesson from fitting judgments of perceived depth. *Journal of Experimental Psychology: General, 121*, 364–381.

Diekhoff, G. (1992). *Statistics for the social and behavioral sciences: Univariate, bivariate, multivariate*. Dubuque, IA: Wm. C. Brown.

Deutsch, D. (1978). Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics, 23*, 227–230.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. Orlando, FL: Academic Press.

Ellingson, T. (1992). Transcription. In H. Myers (Ed.), *Ethnomusicology: An introduction* (pp. 110–152). New York: Norton.

2. Funding for this research was provided by the Natural Sciences and Engineering Research Council of Canada. Emmanuel Bigand, Katy Dunham, Mari Jones, Bill Thompson, and Sandra Trehub provided helpful comments on earlier versions of the manuscript.

Jones, M. R. (1981). Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology, 1* (2), 34–51.
Jones, M. R. (1982). Music as a stimulus for psychological motion: Part II. An expectancy model. *Psychomusicology, 2* (1), 1–13.
Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41*, 621–634.
Jones, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology, 9*, 193–228.
Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459–491.
Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics, 32*, 211–218.
Jones, M. R., Summerell, L., & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *The Quarterly Journal of Experimental Psychology, 39A*, 89–121.
Kallman, H. J. (1982). Octave equivalence as measured by similarity ratings. *Perception & Psychophysics, 32*, 37–49.
Koffka, K. (1935). *Principles of Gestalt psychology*. London: Routledge & Kegan Paul.
Kohler, W. (1947). *Gestalt psychology: An introduction to new concepts of modern psychology*. New York: Liveright.
Krumhansl, C. L. (1995). Music psychology: Influences from music theory. *Music Theory Spectrum, 17*, 53–80.
Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Chicago: University of Chicago Press.
Morton, D. (1980). Thailand. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (Vol. 18, pp. 712–722). London: Macmillan.
Narmour, E. (1989). The "genetic code" of melody: Cognitive structures generated by the implication-realization model. *Contemporary Music Review*, 1989,4, 45–64.
Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model.* Chicago: University of Chicago Press.
Narmour, E. (1992). *The analysis and cognition of melodic complexity: The implication-realization model.* Chicago: University of Chicago Press.
Rosner, B. S., & Meyer, L. B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 317–341). New York: Academic Press.
Rosner, B. S., & Meyer, L. B. (1986). The perceptual roles of melodic process, contour and form. *Music Perception, 4*, 1–40.
Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58*, 75–125.
Schellenberg, E. G., & Trehub, S. E. (1996). Children's discrimination of melodic intervals. *Developmental Psychology, 32*, 1039–1050.
Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*, 109–150.
Schmuckler, M. A. (1990). The performance of global expectancies. *Psychomusicology, 9*, 122–147.
Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
Unyk, A. M., & Carlsen, J. C. (1987). The influence of expectancy on melodic perception. *Psychomusicology, 7*, 3–23.

Appendix

The tables in this appendix provide detailed statistics from the multiple regression analyses. For each predictor variable (or set of predictors) in each model, the squared semipartial correlation (sr^2 or SR^2), F ratio, and p value are reported. The squared semipartial correlation represents the proportion of variance in the outcome variable that is uniquely explained by each predictor (or set of predictors).

TABLE A1
Analyses of Data from Schellenberg (1996)

	sr^2	F	df	p
EXPERIMENT 1 (N = 120)				
Model 1: Implication-Realization (I-R) Model				
model R^2 = .683				
Registral direction	.029	10.12	1, 113	<.005
Intervallic difference	.024	8.22	1, 113	<.005
Registral return	.067	23.63	1, 113	<.0001
Proximity	.089	31.56	1, 113	<.0001
Closure	.036	12.56	1, 113	<.001
Tonality covariate	.015	5.13	1, 113	<.05
Model 2: Revised Model				
model R^2 = .759				
Registral direction–revised	.084	40.29	1, 115	<.0001
Registral return–revised	.052	24.91	1, 115	<.0001
Proximity–revised	.472	225.56	1, 115	<.0001
Tonality covariate	.034	16.30	1, 115	<.0001
Model 3: Principal-Components Model (I-R)				
model R^2 = .669				
Factor 1	.516	179.12	1, 115	<.0001
Factor 2	.077	26.57	1, 115	<.0001
Factor 3	.036	12.22	1, 115	<.001
Tonality covariate	.016	5.51	1, 115	<.05
Model 4: Principal-Components Model (Revised)				
model R^2 = .759				
Factor 1	.214	102.89	1, 116	<.0001
Factor 2	.490	235.45	1, 116	<.0001
Tonality covariate	.035	16.64	1, 116	<.0001
Model 5: Two-Factor Model				
model R^2 = .759				
Pitch proximity	.476	229.15	1, 116	<.0001
Pitch reversal	.207	99.80	1, 116	<.0001
Tonality covariate	.034	16.41	1, 116	<.0001
EXPERIMENT 2 (N = 200)				
Model 1: Implication-Realization (I-R) Model				
model R^2 = .461				
Registral direction	.093	33.63	1, 194	<.0001
Intervallic difference	.017	5.93	1, 194	<.05
Registral return	.035	12.42	1, 194	<.001
Proximity	.071	25.69	1, 194	<.0001
Closure	.016	5.60	1, 194	<.05
Model 2: Revised Model				
model R^2 = .527				
Registral direction–revised	.094	38.90	1, 196	<.0001
Registral return–revised	.016	6.50	1, 196	<.05
Proximity–revised	.358	148.16	1, 196	<.0001

TABLE A1 (CONTINUED)

	sr^2	F	df	p
Model 3: Principal-Components Model (I-R)				
model R^2 = .447				
Factor 1	.311	110.24	1, 196	<.0001
Factor 2	.036	12.94	1, 196	<.0005
Factor 3	.103	36.61	1, 196	<.0001
Model 4: Principal-Components Model (Revised)				
model R^2 = .524				
Factor 1	.144	59.44	1, 197	<.0001
Factor 2	.381	157.53	1, 197	<.0001
Model 5: Two-Factor Model				
model R^2 = .520				
Pitch proximity	.364	149.38	1, 197	<.0001
Pitch reversal	.144	59.26	1, 197	<.0001
EXPERIMENT 3 (N = 132)				
Model 1: Implication-Realization (I-R) Model				
model R^2 = .690				
Registral direction	.010	3.97	1, 125	<.05
Intervallic difference	.065	25.85	1, 125	<.0001
Registral return	.031	12.42	1, 125	<.001
Proximity	.067	26.76	1, 125	<.0001
Closure	.048	19.31	1, 125	<.0001
Tonality covariate	.009	3.26	1, 125	<.1
Model 2: Revised Model				
model R^2 = .755				
Registral direction–revised	.077	39.73	1, 127	<.0001
Registral return–revised	.015	7.75	1, 127	<.01
Proximity–revised	.570	295.10	1, 127	<.0001
Tonality covariate	.010	4.96	1, 127	<.05
Model 3: Principal-Components Model (I-R)				
model R^2 = .688				
Factor 1	.615	250.42	1, 127	<.0001
Factor 2	.043	17.52	1, 127	<.0001
Factor 3	.017	6.86	1, 127	<.01
Tonality covariate	.008	3.41	1, 127	<.1
Model 4: Principal-Components Model (Revised)				
model R^2 = .754				
Factor 1	.122	63.10	1, 128	<.0001
Factor 2	.599	311.29	1, 128	<.0001
Tonality covariate	.010	5.09	1, 128	<.05
Model 5: Two-Factor Model				
model R^2 = .750				
Pitch proximity	.579	296.22	1, 128	<.0001
Pitch reversal	.120	61.41	1, 128	<.0001
Tonality covariate	.010	5.15	1, 128	<.05

TABLE A2

Analyses of Data from Cuddy and Lunney (1995, Appendix) $N = 200$

	sr^2	F	df	p
Model 1: Implication-Realization (I-R) Model				
model $R^2 = .640$				
Registral direction	.008	4.25	1, 191	<.05
Intervallic difference	.022	11.46	1, 191	<.001
Registral return	.026	13.34	1, 191	<.0005
Proximity	.034	17.85	1, 191	<.0001
Closure	.008	3.89	1, 191	<.1
Tonality and height covariates	.362	63.93	3, 191	<.0001
Model 2: Revised Model				
model $R^2 = .725$				
Registral direction–revised	.018	12.56	1, 193	<.0005
Registral return–revised	.025	17.38	1, 193	<.0001
Proximity–revised	.304	213.02	1, 193	<.0001
Tonality and height covariates	.379	88.56	3, 193	<.0001
Model 3: Principal-Components Model (I-R)				
model $R^2 = .630$				
Factor 1	.246	127.96	1, 193	<.0001
Factor 2	.023	11.89	1, 193	<.001
Factor 3	.014	7.03	1, 193	<.01
Tonality and height covariates	.359	62.43	3, 193	<.0001
Model 4: Principal-Components Model (Revised)				
model $R^2 = .721$				
Factor 1	.070	48.12	1, 194	<.0001
Factor 2	.304	210.83	1, 194	<.0001
Tonality and height covariates	.376	86.85	3, 194	<.0001
Model 5: Two-Factor Model				
model $R^2 = .724$				
Pitch proximity	.303	212.41	1, 194	<.0001
Pitch reversal	.066	55.81	1, 194	<.0001
Tonality and height covariates	.378	88.37	3, 194	<.0001