Music Perception
Spring 1997, Vol. 14, No. 3, 263–280

Perceptual Judgments of Triads and Dyads: Assessment of a Psychoacoustic Model

WILLIAM FORDE THOMPSON
York University, Ontario, Canada

RICHARD PARNCUTT
Keele University, Staffordshire, England

In two experiments, goodness-of-fit ratings of pairs of musical elements (triads, dyads, and octave-complex tones) were examined in view of a psychoacoustic model. The model, referred to as the pitch commonality model, evaluates the sharing of fundamental frequencies, overtones, and subharmonic tone sensations between sequential elements and also considers the effects of auditory masking within each element. Two other models were also assessed: a reduced model that considers the sharing of fundamental frequencies alone and the cycle-of-fifths model of key and chord relatedness. In Experiment 1, listeners rated the goodness of fit of 12 octave-complex tones following a major triad, major-third dyad, and perfect-fifth dyad. Multiple regression revealed that pitch commonality provided predictive power beyond that of the reduced model. A regression model based on pitch commonality and the cycle of fifths had a multiple *R* of .92. In Experiment 2, listeners rated how well a triad or dyad followed another triad or dyad. All pairings of the major triad, major-third dyad, and perfect-fifth dyad (pair types) were presented at various transpositions with respect to one another. Multiple regression revealed that pitch commonality again provided predictive power beyond that of the reduced model. A regression model based on pitch commonality, the cycle of fifths, and a preference for trials ending with a triad had a multiple *R* of .84. We discuss the role of psychoacoustic factors and knowledge of chord and key relationships in shaping the perception of harmonic material.

THE analysis of Western music places considerable emphasis on the identification of chords (e.g., Piston, 1987; Salzer, 1952/1962; Shirlaw, 1917/1969). The perceptual and cognitive basis of musical harmony, how-

Address correspondence to William Forde Thompson, Department of Psychology, Atkinson College, York University, 4700 Keele Street, North York, Ontario, M3J 1P3, Canada (e-mail: billt@yorku.ca) or Richard Parncutt, Department of Psychology, Keele University, Staffordshire, ST5 5BG, England.

ever, is not fully understood. Explanations of harmony have considered voice leading and hierarchical structure (e.g., Schenker, 1906/1954, 1935/1979), the importance of simple integer ratios between the fundamental frequencies of chord notes (for a review, see Tenney, 1988), interactions among overtones (Helmholtz, 1863/1954; Plomp & Levelt, 1965), and schemata developed from long-term exposure to music (Krumhansl, Bharucha, & Kessler, 1982; Krumhansl & Kessler, 1982). In this study, we assessed the influence of psychoacoustic factors on goodness-of-fit ratings of pairs of musical elements (triads, dyads, and octave-complex tones), drawing from the psychoacoustic model of harmony proposed by Terhardt (1974) and elaborated by Parncutt (1989; Parncutt & Strasburger, 1994).

Early accounts of harmony were based on Pythagoras's view that preferred combinations of notes have simple integer ratios between their frequency of vibration (Tenney, 1988). As an alternative to this account, Helmholtz (1863/1954) proposed that dissonance arises from physical beating patterns that result when near-coincident frequencies are sounded together. In this account, note combinations with simple integer ratios are consonant because they are characterized by an absence of such beating (but see Schellenberg & Trehub, 1994).

More recent discussions have emphasized sensory and perceptual processes (e.g., Parncutt, 1989; Plomp & Levelt, 1965; Terhardt, 1974, 1977). For example, Plomp and Levelt (1965) linked dissonance to the concept of a critical bandwidth (see also Kameoka & Kuriyagawa, 1969a, 1969b). When spectral components of two notes fall within a critical bandwidth, a sensation of roughness or dissonance may occur. Consonant note combinations are those in which spectral components tend to fall outside the same critical bandwidths.

Auditory masking effects may also influence the perception of complex sonorities (Moore, 1982; Terhardt, Stoll, & Seewann, 1982). Spectral components with different frequencies can mask each other, reducing their audibility. Masking is strongest between components that fall within a critical bandwidth, for example, components separated by a semitone. Although the precise influence of masking on the perception of chords is unknown, it is likely that auditory masking modifies both the dissonance associated with near-coincident frequency components and the salience of individual pitch sensations within a chord.

Finally, Terhardt (1974, 1977) and Parncutt (1989) have argued that subharmonic tone sensations are implicated in the perception of harmonic stimuli, most notably the perception of chord roots. Subharmonic tone sensations are best exemplified in the phenomenon of the missing fundamental, in which the fundamental frequency of a complex periodic wave is physically absent, but is perceived as a *virtual pitch* (Houtsma & Goldstein, 1972; Ritsma, 1967).

Description of the Psychoacoustic Model

Drawing from Terhardt's (1972) theory of pitch perception (see also Terhardt, Stoll, & Seewann, 1982), Parncutt (1989; Parncutt & Strasburger, 1994) proposed a model of how psychoacoustic factors act in combination to influence perceived relationships between complex musical sonorities. The model involves two stages. The first stage estimates the salience of pitches in individual sonorities. The second stage uses the output from the first stage to predict perceived relationships between successive sonorities.[1]

The input to Parncutt's model is a simultaneity of steady-state pure-tone components. Each is assigned a frequency (Appendix, Eq. 1) and an intensity (expressed as sound pressure level or SPL in decibels).[2] Phase is ignored. The model first evaluates the extent of masking among pure-tone components. Components that are entirely masked are subsequently ignored. Partially masked components are assigned appropriately reduced weights in the subsequent analysis. The masking algorithm considers the pure-tone heights of all input frequencies (Eq. 2). The degree to which components mask each other is influenced by their intensity and the distance between them on the pure-tone height scale, measured in critical bandwidths (Eq. 3).[3] The combined effect of several maskers is estimated by adding amplitudes (Eq. 4; see also Terhardt, 1979).

The *audible level* of each pure-tone component is then calculated (Eq. 5).[4,5] Audible levels are scaled by a saturation function (Eq. 6), the output

1. The present implementation of the model is identical to that described by Parncutt and Strasburger (1994), with one exception. In the latter implementation, pitch commonality was estimated by finding the simple correlation between salience vectors. In this case, pitch commonality is estimated by using the formula outlined in Parncutt (1989, eq. 4.29).

2. Participants in the present study were able to adjust the overall loudness to a comfortable level. Therefore, SPL values were estimated on the basis of a sample of listeners.

3. The model assumes that excitation patterns are triangular and symmetric when plotted on the pure-tone-height scale. Models involving asymmetric gradient patterns have been explored by the second author, but they provided no additional predictive power. Excitation patterns for isolated pure tones are asymmetric at higher sound pressure levels, an effect known as upward spread of masking (Moore, 1982).

4. Audible level is the level of a pure-tone component above masked threshold. It is similar to sensation level but applies to individual components. Sensation level generally refers to the level of a whole sound above threshold.

5. This procedure neglects the threshold of hearing and as such is only appropriate for midrange (i.e., musically typical) frequencies and levels (the sonorities of the present experiments conformed to this restriction).

of which is a value between 0 (inaudible) and 1 (clearly audible). This value, called the *audibility* of the pure-tone component, is a measure of the likelihood both that the component will be separately heard and that it will influence the perception of any complex tone of which it is a partial.

Next, harmonic patterns among audible pure-tone components are identified by a pattern-recognition procedure. The pattern-recognition template consists of the first 10 harmonics of a complex tone (Eq. 7). All pitches are expressed relative to the chromatic scale, that is, rounded to the nearest semitone. Components of the template are weighted such that higher harmonics play a less important role than lower harmonics in the matching algorithm (Eq. 8).[6]

The template is shifted by semitone steps through the entire pitch range, searching for harmonic pitch patterns embedded among the audible pure-tone components. When a harmonic pattern is matched, a *complex-tone sensation* is created, whose ("virtual") pitch corresponds to the fundamental frequency of the pattern.

The *audibility* of the complex tone is estimated by evaluating the goodness of fit between the template and the audible sound spectrum (Eq. 9). The audibility of each complex tone is compared with the audibility of the pure-tone component at the fundamental frequency of that complex, and the greater of the two values is taken as the *salience* of the pitch category (Eq. 10).[7] A salience value is obtained for each pitch category in the chromatic scale, resulting in a vector of pitch salience values. Judgments of how well one musical element follows another (e.g., a triad followed by another triad) are predicted to correspond to *pitch commonality,* a measure of the overlap between the pitch salience vectors of the two elements (Eq. 11).

Pitch commonality was used to predict goodness-of-fit ratings of pairs of musical elements.[8] In Experiment 1, listeners rated the goodness of fit of octave-complex tones following a triad or dyad. In Experiment 2, listeners rated how well a triad or dyad followed another triad or dyad. Ratings were compared with corresponding values of pitch commonality. Predictions for Experiment 1 were pitch commonalities between dyads or triads and octave-complex tones. Predictions for Experiment 2 were pitch commonalities between pairs of dyads or triads.

6. This weighting reflects the relative amplitudes of the complex tones in the triads and dyads presented in Experiments 1 and 2. Other formulas for weighting components of the template provided no additional predictive power in Parncutt's (1989) assessment of the model.

7. In previously published versions of the model, salience is normalized by multiplying audibility by a scaling factor that depends on the total number of tones or pitches simultaneously perceived in the sonority (its "multiplicity"). We omitted this procedure here as it has no effect on the estimation of pitch commonality.

8. Experimental data were collected by the first author at the Department of Psychology, Queen's University, under the supervision of Lola L. Cuddy.

The Cycle-of-Fifths Model

We expected that our experimental data would reflect both psychoacoustic factors and learned schemata. Thus, in addition to the factors outlined above, we evaluated another predictor: *the cycle of fifths*. The cycle of fifths is well known as a model of key relatedness, but it is used here to model the relatedness between pairs of musical elements.

As a model of key relatedness, the cycle of fifths maps keys in a circular arrangement such that the tonic notes (i.e., the first scale degree, or *doh*) of adjacent major keys are separated by a perfect fifth. In particular, the fifth scale degree of each key (i.e., *sol*) is the first scale degree of the adjacent key in a clockwise direction on the cycle (e.g., A major and E major). Key distance is defined by the number of steps on the cycle. Neighboring major keys share all but one scale degree, and the number of scale degrees shared between keys decreases as key distance increases. A number of studies have supported or elaborated on the cycle of fifths as a psychological model of key relatedness (e.g., Cuddy & Thompson, 1992; Krumhansl & Kessler, 1982; Thompson, 1993; Thompson & Cuddy, 1989, 1992, 1997; for an overview, see Krumhansl, 1990).

In this study, we evaluated the extent to which pitch commonality, together with the cycle-of-fifths model, can account for goodness-of-fit judgments of triads, dyads, and octave-complex tones. Listeners judged how well octave-complex tones followed triads and dyads (Experiment 1) and how well triads and dyads followed other triads and dyads (Experiment 2). The predictive power of pitch commonality was contrasted with the predictive power of a simpler model (the *reduced* model) that considered only note commonality. Finally, we evaluated predictions based on the cycle-of-fifths model.

Experiment 1

In Experiment 1, listeners were presented a triad or dyad followed by each of 12 probe tones: one for each pitch class in the equal-tempered chromatic scale. Listeners rated how well each probe tone fit with the triad or dyad. Probe tones were octave-complex (circular) tones, similar in construction to Shepard tones (Shepard, 1964) but with additional partials. Probe-tone profiles (i.e., the set of 12 mean probe-tone ratings) were obtained for a major triad, a major-third dyad, and a perfect-fifth dyad. Probe-tone profiles were compared with predictions based on pitch commonality and the cycle-of-fifths model. A reduced model based on note commonality was also assessed. By comparing the results for the full model with the results for the reduced model, we were able to evaluate the importance of overtones and subharmonics.

METHOD

Listeners

Fifteen listeners, ranging in age from 18 to 32 years, participated in this experiment. Participants were at a moderate level of musical training, having had a minimum of 5 years of formal training on a musical instrument. All participants reported having normal hearing and were naive concerning the purposes of the experiment.

Apparatus and Stimuli

Tones were produced by a DMX-1000 real-time digital synthesizer controlled by a PDP 11/23 computer. Presentations consisted of an initial simultaneous set of either two or three tones, 350 ms in duration, followed by a 500-ms pause and then a probe tone of 1-s duration. All tones had a rise and decay time of 22 ms each. Each of the initial set of tones had five partials, with the amplitude of each partial inversely proportional to the partial number. Fundamental frequencies were taken from the equal-tempered pitch set, and amplitudes of fundamentals were determined by Fletcher-Munson equal-loudness contours.

Probe tones were similar in construction to Shepard tones (Shepard, 1964), but involved additional partials. They were produced by simultaneously sounding nine octave-equivalents of a single complex tone, within the range from A_0 to A_9. Each of the nine complex tones had the same construction as the tones used in the initial musical element (i.e., five partials). The amplitudes of the nine complex tones were adjusted so as to approximate a Gaussian envelope over the entire set. The peak of this envelope was held constant for all probe tones. Responses were made via a computer terminal in a soundproof booth, and sequences were heard by participants through Sennheiser HD-480 headphones.

Procedure and Conditions

On each trial, participants were presented with a simultaneous set of two or three tones, followed by a rest and then a probe tone. The simultaneous set of two tones was either a major-third dyad or a perfect-fifth dyad. The simultaneous set of three tones was a major triad in root position. All twelve notes of the chromatic scale were used as probe tones for each of the three musical elements. Participants were presented with each combination of musical element and probe tone once. Each presentation was randomly transposed from trial to trial with the condition that the lowest tone in the triad or dyad was at one of the equal-tempered pitches between G_3 and C_4. The order of trials was completely scrambled for each participant. Participants were asked to rate on a scale of 1 to 7 how well the probe tone fit the initial musical element. Participants were informed that there were no right or wrong answers and that they should try to use the full range of the response scale.

RESULTS AND DISCUSSION

Mean probe-tone ratings for the perfect-fifth dyad, major-third dyad, and major triad are shown in Table 1. For clarity, means are displayed as though the lowest tone of the musical element were always C, and probe tones will henceforth be referenced by their corresponding note name in the table. Each profile of 12 mean ratings was analyzed by using multiple comparisons in a one-way analysis of variance (ANOVA). Next, correla-

TABLE 1
Mean Ratings of Probe Tones Following a Major Triad, a Major-Third Dyad, and a Perfect-Fifth Dyad

	Position of Probe Tone											
	C	C♯	D	D♯	E	F	F♯	G	A♭	A	B♭	B
Triad	6.3	2.5	3.5	2.9	4.7	4.7	2.9	5.8	3.5	4.1	4.3	3.3
Third	6.3	3.3	3.6	3.5	5.3	3.8	3.1	5.0	3.3	4.9	3.1	3.6
Fifth	6.3	2.8	4.7	3.9	4.1	3.9	2.3	6.1	3.9	4.1	3.9	3.5

Ratings are displayed as though the lowest note of the triad or dyad were always C.

tion and regression were conducted to evaluate predictions based on pitch commonality, a reduced model based on fundamental frequencies alone, and predictions based on the cycle of fifths.

For the triad, probe tones with the same pitch class as fundamental frequencies in the presentation (C, E and G in Table 1) were assigned significantly higher ratings than the other nine probe tones [$F(1, 14) = 34.37$, $p < .001$]. Second, of these probe tones, E (the mediant) was assigned lower ratings than were C and G [$F(1, 14) = 20.40$, $p < .001$]. One interpretation of this effect is that the mediant was subjected to a greater degree of masking than were the other triadic tones (i.e., it was masked both from above by the dominant and from below by the tonic). Among the remaining nine probe tones, the probe tone at F was given the highest average rating [$F(1, 14) = 18.77$, $p < .001$]. This finding has several possible interpretations. It may reflect a sensitivity to key relatedness (the key of F major is one step on the cycle of fifths from the key of C major), sensitivity to typical harmonic motion (C major chords are often followed by F chords), or sensitivity to subharmonic tone sensations (F has the pitch class of a near subharmonic of the lowest note of the triad, C). The probe tone at B♭ was also rated relatively highly following the major triad, even though this probe tone does not represent a near harmonic or subharmonic of any note in the major triad. However, B♭ represents the tonic of a related key. In addition, triads are often associated with the flattened seventh in harmony (e.g., as a dominant seventh chord).

For the major-third dyad, probe tones with the same pitch class as fundamental frequencies in the presentation (C and E in the table) were rated higher than other probe tones [$F(1, 14) = 68.42$, $p < .001$]. Second, of the remaining 10 probe tones, G was rated the highest [$F(1, 14) = 8.50$, $p < .05$]. A sensory interpretation is that G was physically present as the third harmonic of the lower tone in the interval. Alternatively, the dyad may have been interpreted as the lower two notes of an incomplete major triad. Third, of the remaining nine probe tones, the probe tone at A was rated significantly higher than the others [$F(1, 14) = 10.31$, $p < .01$]. A sensory interpretation is that this probe tone has the same pitch class as a subharmonic of the upper note in the interval (E in the table). Another possibility is that the dyad was interpreted as the upper two notes of an incomplete minor triad. The probe tone at F was not rated highly, even though this probe tone represents the pitch class of a subharmonic of the lower note of the interval (C).

For the perfect-fifth dyad, probe tones with the same pitch class as fundamental frequencies in the presentation (C and G in the table) were again rated higher than other probe tones [$F(1, 14) = 53.00$, $p < .001$], and these tones were not rated significantly differently from each other. Of the remaining 10 probe tones, the probe tone at D was rated significantly higher than others [$F(1, 14) = 8.29$, $p < .05$]. A sensory interpretation of this finding is that D was physically present as the third harmonic of the complex tone whose fundamental was at G. The probe tone at F was not rated highly, even though this probe tone represents the pitch class of a subharmonic of the lower note of the interval (C).

Correlation and Regression

The second analysis involved correlation and regression. We first observed that Krumhansl and Kessler's (1982) major triad profile was highly correlated with our major triad profile ($r = .92$, $p < .01$), our major-third dyad profile ($r = .86$, $p < .01$), and our perfect-fifth dyad profile ($r = .83$, $p < .01$).

Next, we examined the complete set of 36 mean ratings: 12 mean ratings each for the major triad, major-third dyad, and perfect-fifth dyad. Three predictors were evaluated: pitch commonality, the reduced model, and the cycle of fifths. Predictions based on pitch commonality corresponded to the overlap of pitch salience values estimated for the triad or dyad and the octave-complex tones (Appendix, Eq. 11). The reduced model predicted that probe tones would be assigned high ratings only if they had the same pitch class as a fundamental frequency in the triad or dyad. For example, for the triad, the reduced model predicted that probe tones representing the three triad notes (C, E, and G in the table) would be assigned high ratings and other probe tones would be assigned low ratings. The cycle-of-fifths predictor was created by coding each probe tone as the number of steps around the cycle of fifths between it and the lowest note of the triad or dyad (C in the table). For example, if a C major triad was followed by a probe tone at E, that probe tone would be assigned a predicted value of four, because E is four steps removed from C on the cycle of fifths.

TABLE 2
Correlation Matrix of Mean Ratings and Three Predictors, and Multiple Regression Results, for Experiment 1

	Pearson Correlation Results		
	Ratings	Pitch Commonality	Reduced Model
Pitch commonality	0.84		
Reduced model	0.80	0.96	
Cycle of fifths	–0.73	–0.49	–0.43

	Multiple Regression Results				
Variable	Coefficient	*SD* error	*SD* coeff	*t*	*P*
Constant	4.34	0.20	0.00	22.14	<.001
Pitch commonality	2.26	0.28	0.63	8.08	<.001
Cycle of fifths	–0.26	0.05	-0.43	–5.41	<.001

Dependent variable = mean rating, $N = 36$, multiple $R = 0.919$, multiple $R^2 = 0.845$.

Table 2 provides the correlation matrix for the set of mean ratings and the three predictors and also displays the multiple regression results. The correlation analysis revealed that all three predictors were highly correlated with the set of mean ratings and were significantly correlated with each other. The correlation between mean probe-tone ratings and predictions based on pitch commonality was higher than the correlation between mean probe-tone ratings and predictions based on the reduced model. Moreover, when mean probe-tone ratings were regressed on these predictors, pitch commonality had predictive power significantly beyond that provided by the reduced model ($t = 2.88$, $p < .01$).

The reduced model had no significant predictive power beyond that provided by the other predictors ($t = .166$, not significant) and was therefore removed from the final regression model. Thus, the final regression model included pitch commonality and the cycle of fifths. The multiple correlation for these two predictors combined ($R = .92$) was highly significant, suggesting that the psychoacoustic predictions of pitch commonality, with predictions based on the cycle of fifths, provide an excellent account of the variance in mean ratings.

Because Experiment 1 involved presenting triads and dyads in isolation, psychoacoustic factors may have been more salient than they would have been in a broader musical context (Krumhansl, 1983). Thus, as a further examination of the importance of psychoacoustic factors, the second experiment involved obtaining judgments of pairs of triads and dyads.

Experiment 2

In Experiment 2, all possible pairings of major triads, major-third dyads, and perfect-fifth dyads were presented to listeners at various transpositions with respect to each other. Participants judged how well the second musical element followed the first. As in Experiment 1, the results were evaluated in view of pitch commonality, the reduced model, and the cycle-of-fifths model.

METHOD

Participants

Fifteen listeners, ranging in age from 19 to 28 years, participated in this experiment. All participants were highly trained in music, having taken at least 10 years of instruction on a musical instrument and 2 years of music theory. All participants reported having normal hearing and were naive concerning the purposes of the experiment.

Apparatus and Stimuli

The apparatus and testing environment were the same as in Experiment 1. Presentations consisted of an initial simultaneous set of either two or three tones, 350 ms in duration, followed by a 500-ms pause and then another simultaneous set of either two or three tones 350 ms in duration. Each tone was constructed in the same manner as those used in Experiment 1.

Procedure and Conditions

On each trial, participants were presented with a triad or dyad, followed by a rest, and then another triad or dyad. Each musical element was one of the following musical elements: a major triad; a dyad of a major third; and a dyad of a perfect fifth. All nine possible pairs of musical elements were used as stimuli. The first musical element was randomly transposed to one of six places such that the lowest tone of the triad or dyad was at $B\flat_3$, B_3, C_4, G_4, $A\flat_4$, or A_4. For each of the nine pair types, the second musical element was transposed with respect to the first musical element to each of the 12 chromatic tones, randomly and without replacement. If the lowest tone of the first musical element was set at $B\flat_3$, B_3, or C_4, transposition brought the lowest tone of the second element either to the same pitch height or to a higher pitch height. If the lowest tone of the first musical element was set at G_4, $A\flat_4$, or A_4, transposition brought the lowest tone of the second element either to the same pitch height or to a lower pitch height. This procedure was used to emphasize the effects of pitch-class relationships, rather than pitch proximity. The order of trials was scrambled for each participant. Participants were asked to rate on a scale of 1 to 7 how well the second musical element fit the initial musical element. Participants were informed that there were no right or wrong answers and that they should try to use the full range of the response scale.

RESULTS AND DISCUSSION

Mean ratings for each pair type are shown in Table 3. The table is displayed with the root of the first musical element set to C. An examination

TABLE 3
Mean Ratings of Nine Pairs of Musical Elements Presented at 12 Different Transpositions with Respect to One Another (Assuming Octave Equivalence)

	Lowest Note of Second Musical Element											
	C	C♯	D	D♯	E	F	F♯	G	A♭	A	B♭	B
Triad-triad	6.4	4.1	4.5	4.9	4.1	6.1	3.7	5.9	5.0	4.3	4.0	3.3
Triad-third	6.3	3.5	4.0	3.8	3.4	4.0	3.3	3.7	3.3	3.0	3.7	2.7
Triad-fifth	5.7	3.3	3.7	5.1	3.1	5.1	2.9	5.0	3.2	4.0	3.9	2.5
Third-triad	6.7	3.9	4.9	4.1	4.9	5.8	3.7	5.4	4.9	5.7	4.5	3.5
Third-third	6.1	3.5	4.7	4.3	4.0	4.3	3.5	4.9	4.5	4.6	3.0	4.0
Third-fifth	6.4	3.9	3.6	3.1	4.9	4.7	3.3	4.7	4.0	5.2	4.5	3.0
Fifth-triad	6.3	4.2	4.1	4.5	3.9	5.3	3.9	5.7	4.3	4.6	4.2	3.5
Fifth-third	6.3	4.2	4.5	4.4	3.1	5.0	3.4	5.2	4.0	3.3	4.3	3.3
Fifth-fifth	6.1	3.4	4.9	4.5	4.1	6.1	4.1	5.8	3.9	4.1	4.5	3.9
Mean	6.2	3.8	4.3	4.3	4.0	5.2	3.5	5.1	4.1	4.3	4.0	3.3

Ratings are displayed as though the lowest note of the first musical element were always C.

of mean ratings across all pair types (bottom row of table) reveals a correspondence between ratings and relationships between keys as described by the cycle of fifths. Collapsed across all pair types, pairs having the same lowest note (for the triad, the root; for the dyads, the root of an implied chord) were given higher ratings than those whose lowest note moved to G, F, D, and B♭ (one or two steps on the cycle of fifths) [$F(1, 14) = 90.30$, $p < .001$]. Second, pairs whose lowest note moved to G or F (one step on the cycle of fifths) were given higher ratings than pairs whose lowest note moved to D or B♭ (two steps on the cycle of fifths) [$F(1, 14) = 65.23$, $p < .001$]. These differences parallel judgments of key relatedness, suggesting a common or similar underlying mental representation. Psychoacoustic factors can also explain the findings because chords whose roots are adjacent on the cycle of fifths tend to share more harmonic and subharmonic pitches than chords whose roots are not adjacent on the cycle of fifths.

Correlation and Regression

Correlation and regression were used to evaluate the influence of psychoacoustic factors on judgments. We again evaluated the predictors assessed in Experiment 1. The criterion variable consisted of the set of 108 mean ratings (12 transpositions for each of nine pair types). Predictions based on pitch commonality were established by comparing pitch salience values of the first and second elements for every pair and transposition and then averaging the set of predicted values for rising motion with the set of predicted values for falling motion. The predictions of the reduced model were established by coding the number of shared notes (i.e., fundamental frequencies) between the two elements. For example, a C major triad and a G major triad have one note in common, whereas a C major triad and a D major triad have no notes in common. Predictions based on the cycle of fifths considered the number of steps on the cycle of fifths between the lowest notes of the two elements. For example, a C major triad followed by an E major triad was assigned a value of “4,” because C and E are four steps removed on the cycle of fifths.

Table 4 displays the correlation matrix for mean ratings obtained in Experiment 2 and three predictors: pitch commonality, the reduced model, and the cycle of fifths. The correlation between mean ratings and predictions based on pitch commonality was higher than the correlation between mean ratings and predictions based on the reduced model. Moreover, when the set of mean ratings was regressed on these two predictors, pitch commonality had predictive power significantly beyond that of the reduced model ($t = 3.43$, $p < .001$). However, the reduced model had no predictive power significantly beyond that of pitch commonality ($t = -1.10$, $p = .27$).

We next assessed the combined predictive power of the predictors by using multiple regression. Pitch commonality and the cycle of fifths were included in the regression model, but the reduced model was excluded because it provided no significant predictive power beyond that provided by the other predictors. These two predictors—pitch commonality and the cycle of fifths—accounted for 65% of the variance in mean ratings ($R = .81$, $p < .001$).

TABLE 4
Correlation Matrix of Mean Ratings and Three Predictors, and Multiple Regression Results, for Experiment 2

	Pearson Correlation Results		
	Ratings	Pitch Commonality	Reduced Model
Pitch commonality	0.72		
Reduced model	0.69	0.98	
Cycle of fifths	–0.73	–0.62	– 0.55

	Multiple Regression Results				
Variable	**Coefficient**	**STD error**	**STD coeff**	***t***	***p***
Constant	4.42	0.190	0.00	23.16	<.001
Pitch commonality	0.53	0.089	0.41	5.96	<.001
Cycle of fifths	–0.26	0.037	–0.48	–7.03	<.001
Movement	0.46	0.110	0.23	4.26	<.001

Dependent variable = mean rating, $N = 108$, $R = 0.838$, $R^2 = 0.703$

Another predictor (preferred movement) was then added to account for overall differences in mean ratings for different pair types. In particular, an examination of residuals suggested that mean ratings were higher for pair types in which the second element was a triad. Although the reasons for this tendency cannot be determined from the present data, it is notable that triadic endings are more common than dyadic endings in Western tonal harmony. It is possible that movement to a triad generally conveyed greater harmonic resolution than movement to a dyad. For this predictor, conditions involving movement to a triad were coded as 1 and other conditions were coded as 0. The results of the multiple regression analysis are displayed in Table 4.

The multiple correlation for the three predictors combined ($R = .84$) was highly significant and indicates that the combined model, which includes pitch commonality, the cycle of fifths, and preferred movement, accounted for 70% of the variance in mean ratings.

Finally, we examined whether the set of mean ratings obtained in Experiment 2 could be predicted from mean probe-tone ratings obtained in Experiment 1. For each pair type, the set of mean probe-tone ratings for each element was taken from Experiment 1. We then obtained the correlation between the two sets of probe-tone ratings at each of 12 possible transpositions, one for each of the 12 transposition conditions in Experiment 2. The resultant 12 correlation coefficients were taken as the predicted ratings for that pair type. To predict ratings of pair types having the same root, the first probe-tone rating of Element 1 was matched with the first probe-tone rating of Element 2, and so on. If both elements were the same (e.g., a C major triad followed by another C major triad), the condition was therefore assigned a (correlation) value of 1.0. In order to predict ratings of pair types whose roots were not the same, probe-tone ratings of the two musical elements were transposed with respect to one another before the correlation coefficient was calculated. For instance, to predict the rating of a C major triad followed by a D major triad, probe-tone ratings for two triads were compared such that the first probe-tone rating for the D major triad was matched with the third probe-tone rating of the C major triad. The analysis revealed a fair correspondence between mean ratings of pairs of elements (Experiment 2) and predictions based on probe-tone ratings obtained in Experiment 1. The correlation between predicted and actual mean ratings was .79 ($p < .001$).

Discussion

The results of this investigation reveal a strong relationship between pitch commonality and judgments of triads and dyads. In Experiment 1, significant differences between mean probe-tone ratings of different experimental conditions were interpretable in terms of influences by fundamental frequencies, overtones, and in some cases, subharmonic tone sensations. Further analyses involving multiple regression suggested that the variance in mean ratings of the 36 different conditions could be largely predicted from a combined model that included pitch commonality and a predictor based on the cycle of fifths (multiple $R = .92$). The predictive power of pitch commonality suggests that judgments were partially based on an extraction of the pitch content of the stimulus materials and a subsequent assessment of pitch commonality between the triad or dyad and the probe tone. The significance of the cycle of fifths suggests that listeners also used knowledge of conventional root movement in assigning probe-tone ratings. Alternatively, ratings may have reflected the suitability of probe tones as tonal centers of the implied key.

In Experiment 2, mean ratings of pairs of elements (major triads, major-third dyads, and perfect-fifth dyads) were well modeled by a combined model that included pitch commonality, the cycle of fifths, and preferred movement (multiple $R = .84$). The significance of pitch commonality suggests that ratings were partially based on an extraction of the pitch content of each triad or dyad, and an assessment of the pitch commonality between the two elements. The significance of the cycle-of-fifths predictor suggests that listeners were familiar with common harmonic progressions. Moreover, the result suggests that perceived relationships between triads or dyads are congruent with perceived relationships between keys. The cycle of fifths not only provides a model for the perception of key movement (Cuddy & Thompson, 1992; Thompson & Cuddy, 1989; 1992) but may also guide perceptual judgments of chords and chord relationships. Finally, the preference for trials ending with a triad may reflect the knowledge that, in Western tonal harmony, triadic endings are more common than dyadic endings.

For both experiments, pitch commonality provided predictive power significantly beyond a reduced model based merely on note commonality. This finding suggests that the combined effect of overtones, masking, and subharmonic tone sensations significantly contributed to judgments. A comparison of correlation values in Tables 2 and 4 suggests that including consideration of overtones, masking, and subharmonic tone sensations increased the variance accounted for by 7% in Experiment 1 and 5% in Experiment 2. Thus, while most of the predictive power of pitch commonality derives from its estimation of the commonality of fundamental frequencies, the other components of the pitch commonality model cannot be entirely disregarded without a significant loss in predictive power.

The proportion of variance in mean ratings accounted for by pitch commonality alone was lower in Experiment 2 ($r^2 = .52$, Table 4) than in Experiment 1 (r^2 = .71, Table 2). Possibly, the greater number of conditions assessed in Experiment 2 (108 conditions versus just 36 conditions in Experiment 1) allowed for a greater number of potential sources of variance. The addition of these unidentified sources of variance, in turn, could account for the general reduction in the proportion of variance attributable to the specific influences accounted for by pitch commonality. Another interpretation is that when richer musical contexts are evaluated in psychological studies, new processes may be instantiated, reducing the relative importance of certain psychoacoustic influences. For example, in longer harmonic sequences, individual voices or melodic lines may be segregated into distinct perceptual streams. This process of directing attention to horizontal structure, in turn, may modify the apparent dissonance associated with interactions among frequency components within a chord (Wright & Bregman, 1987).[9]

References

Cuddy, L. L., & Thompson, W. F. (1992). Asymmetry of perceived key movement in chorale sequences: Converging evidence from a probe-tone analysis. *Psychological Research, 54,* 51–59.

Helmholtz, H. L. F. von (1954). On the sensation of tone as a physiological basis for the theory of music (A. J. Ellis, Ed. and Trans). New York: Dover. (Original work published 1863)

Houtsma, A. J. M., & Goldstein, J. L (1972). The central origin of the pitch of complex tones: Evidence from musical interval recognition. *Journal of the Acoustical Society of America, 51,* 520–529.

Kameoka, A., & Kuriyagawa, M. (1969a). Consonance theory, Part I: Consonance of dyads. *Journal of the Acoustical Society of America, 45,* 1451–1459.

Kameoka, A., & Kuriyagawa, M. (1969b). Consonance theory, Part II: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America, 45,* 1460–1469.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception. 1,* 28–62.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* Oxford, England: Oxford University Press.

Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 24–36.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334–368.

Moore, B. C. J. (1982). *An introduction to the psychology of hearing* (2nd ed.). New York: Academic Press.

Moore, B. C. J., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America, 74*(3), 750–753.

9. This research was funded by a research grant from the Natural Sciences and Engineering Research Council of Canada, awarded to the first author. Helpful comments were provided by Lola L. Cuddy, Adrian Houtsma, Brian Moore, and Ernst Terhardt.

Parncutt, R. (1989). *Harmony: A psychoacoustic approach*. Berlin: Springer.
Parncutt, R., & Strasburger, H. (1994). Applying psychoacoustics in composition: "Harmonic" progressions of "nonharmonic" sonorities. *Perspectives of New Music, 32*(2), 88–129.
Piston, W. (1987). *Harmony* (5th ed., revised and expanded by Mark DeVoto). New York: W. W. Norton & Company. (First edition published 1941)
Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38,* 548–560.
Ritsma, R. J. (1967). Frequencies dominant in the perception of the pitch of complex tones. *Journal of the Acoustical Society of America, 42,* 191–198.
Salzer, F. (1962). *Structural hearing* (Vol. 1). New York: Dover publications. (Original work published 1952)
Schenker, H. (1954). *Harmony* (E. M. Borgese, Trans.). Chicago: University of Chicago Press. (Original work published 1906)
Schenker, H. (1979). *Free composition* (E. Oster, Trans.). New York: Longman. (Original work published 1935)
Schellenberg, E. G., & Trehub, S. E. (1994). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin & Review, 1*(2), 191–201.
Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2346–2353.
Shirlaw, M. (1969). *The theory of harmony.* New York: De Capo Press. (Original work published 1917).
Tenney, J. (1988). *A history of 'consonance' and 'dissonance.'* New York: Excelsior Music Publishing Company.
Terhardt, E. (1972). Zur Tonhöhenwahrnehmung von Klängen. *Acustica, 26,* 173–199.
Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55,* 1061–1069.
Terhardt, E. (1977). The two-component theory of musical consonance. In E. F. Evans, J. P. Wilson, eds. *Psychophysics and the physiology of hearing*. London: Academic Press.
Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research, 1,* 155–182.
Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America, 71*(3), 679–688.
Thompson, W. F. (1993). Modeling perceived relationships between melody, harmony, and key. *Perception & Psychophysics, 53,* 13–24.
Thompson, W. F., & Cuddy, L. L. (1989). Sensitivity to key change in Chorale sequences: A comparison of single voices and four-voice harmony. *Music Perception, 7,* 151–168.
Thompson, W. F., & Cuddy, L. L. (1992). Perceived key movement in four-voice harmony and single voices. *Music Perception, 9,* 427–438.
Thompson, W. F., & Cuddy, L. L. (1997). Music performance and the perception of key. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 116–135.
Wright, J. K., & Bregman, A. S. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review, 2,* 63–92.

Appendix: Description of the Model

The model is a revised version of that presented in Parncutt (1989). Each sonority is input to the model as a vector of the *sound pressure level* (SPL, in decibels) of each pure-tone component as a function of its *pitch category* P (in equally tempered semitones above 16.35 Hz, or the note C_0). The pitch category of middle C (C_4, 262 Hz) is $P = 4 \times 12 = 48$, of $C\sharp_4$ is 49, and so on.

The frequency of each pure-tone component in hertz is as follows:

$$f = 440 \times 2^{(P-57)/12} \quad (1)$$

and its *pure-tone height* H_p is given by Moore and Glasberg's (1983) formulation of ERB rate (where ERB stands for equivalent rectangular bandwidth):

$$H_p(f) = 11.17 \log_e \left(\frac{f + 312}{f + 14675}\right) + 43 \quad (2)$$

The degree to which a pure-tone component of frequency f' and SPL(f') masks another component f is given by a partial masking level ml(f,f'):

$$\mathrm{ml}(f,f') = \mathrm{SPL}(f') - 18\,\left|H_p(f') - H_p(f)\right| \quad (3)$$

where 18 is the gradient of the masking pattern in decibels per critical band and the vertical bars denote absolute value.

The *overall masking level* ML (in decibels) at a given frequency f due to several maskers of frequencies f' is estimated by adding amplitudes:

$$\mathrm{ML}(f) = \max\left\{20 \log_{10} \sum_{P \neq P'} 10^{\mathrm{ml}(f,f')/20};\ 0\right\} \quad (4)$$

The *audible level* AL (in decibels) of each pure-tone component is defined as its level above masked threshold:

$$\mathrm{AL}(P) = \max\{\mathrm{SPL}(P) - \mathrm{ML}(P);\ 0\} \quad (5)$$

The *audibility* A_p of each pure-tone component is assumed to saturate with increasing level:

$$A_p(P) = 1 - \exp\left\{\frac{-\mathrm{AL}(P)}{15}\right\} \quad (6)$$

The recognition of harmonic pitch patterns among the audible pure-tone components of a sonority is simulated by matching them against a template of 10 harmonic components. Each component is labeled according to its harmonic number n = 1 through 10. The pitch of each element is given in rounded semitones by its pitch category P_n relative to the lowest element P_1:

$$P_n = P_1 + \mathrm{int}\{12 \log_2(n) + 0.5\} \quad (7)$$

where int denotes integer part.

The components are weighted relative to each other according to the following formula:

$$W_n = 1/n \quad (8)$$

For each sonority, the template is shifted in steps of one semitone through the entire pitch range. In each position, *complex-tone audibility* (A_c) is given by:

$$A_c(P_1) = \frac{1}{3}\left(\sum_n \sqrt{W_n A_p(P_n)}\right)^2 \quad (9)$$

The overall p*itch salience* (S) at pitch category P is given by:

$$S(P) = \max\{A_p(P);\ A_c(P)\} \quad (10)$$

Finally, the *pitch commonality* (C) of two successive sonorities is calculated from their pitch salience vectors S_1 and S_2 according to the following formula:

$$C = \frac{\sum_P \sqrt{S_1(P) S_2(P)}}{\sqrt{\sum_P S_1(P) \sum_P S_2(P)}} \tag{11}$$