Musical and Temporal Influences on Key Discovery

HELEN BROWN
Purdue University

DAVID BUTLER & MARI RIESS JONES
The Ohio State University

The intervallic rivalry model of key identification is outlined and evaluated in two experiments that use a completion judgment task. Experiment 1 replicates an earlier experiment by Cuddy and Badertscher (1987), in which the rare-interval hypothesis of the intervallic rivalry model was considered. In the present study, listeners with different levels of musical training rated probe tones in the context of three different melodic patterns: arpeggiated major triads, ascending major scales, and arpeggiated diminished triads. Results of Experiment 1 indicated that in both the C major triadic and the C major scalar contexts, listeners gave higher completion ratings to all three probes that were members of the presented C major triad than to the other probes, with the exception of F, thereby producing a jagged (multipeaked) profile. For the diminished triadic context, listeners rated the single probe C, that which corresponds to the tonal center in major mode for that group of three tones, as the best completion. Experiment 2 tested the temporal-order hypothesis of the intervallic rivalry model by reordering tones in all three contexts. Again jagged tone profiles appeared with major triadic and major scalar contexts, although in the former the tone F, a perfect fifth below the root of the presented C major triad, received the best completion rating. A single-peaked function appeared with probes in the diminished triadic context, where the major-mode tonic garnered the highest rating found in all conditions of both experiments. Data are interpreted as support for both the rare-interval hypothesis and the temporal-order hypothesis derived from the intervallic rivalry model of key discovery. Complementary findings consistent with the tonal hierarchy model are also discussed.

Introduction

When we listen to music composed in a familiar idiom, we hear coherence in that music. When we listen to music of the Western tonal idiom,
much of that coherence comes from our having developed a listener's vocabulary of relationships among various tones within a special tonal system, one that depends in part on the listener's identification of key. Exactly how a listener quickly and accurately arrives at key identification, however, remains a mystery.

A model that has been proposed to explain key identification within the tonal idiom is the intervallic rivalry model (Brown, 1988; Brown & Butler, 1981; Butler 1989, 1990a; Butler & Brown, 1984). This model describes tonal information that listeners use to discover, in real time, a key or tonal center in music. It rests on three related hypotheses: (1) Primacy—In evaluating rival candidates for the tonic, listeners are biased to assume that the first tone in a musical event is the tonal center, until a better candidate replaces it. This is the primacy hypothesis. (2) Rare intervals—In finally determining a tonal center, listeners rely more on rarer than on common intervals among pitches, those that unambiguously correlate with a single diatonic set, because they provide the more reliable key information. This is the rare-interval hypothesis. (3) Temporal order—In relying on rarer intervals, listeners are more accurate in detecting a correlation between a rare interval and the key of the musical event containing it when the pitches that outline the rare interval appear in a temporal order implying goal-oriented harmonic motions commonly encountered in tonal music. This is the temporal-order hypothesis. The primacy hypothesis is straightforward, suggesting that listeners tend to rely on the first tone they hear as a cognitive anchor (keynote) until it is replaced by a better anchor on the basis of additional musical evidence. The core of the intervallic rivalry model resides in interaction between rare intervals and their temporal orderings in music when heard. These features of the model are defined and illustrated as background to the present study.

![Diagram of major diatonic intervals](image)

**Fig. 1.** Intervals of the diatonic set, arranged in scalar order. The rarer interval within each category of intervals is shown with filled noteheads. Part (a) contains two minor seconds and five major seconds, (b) contains three major thirds and four minor thirds, and (c) contains one augmented fourth and six perfect fourths. The bar graph at the right (d) depicts the interval classes according to number of intervals per class, and as percentages of the total of 21 intervals in the major pitch set.
Musical and Temporal Influences

The rare-interval hypothesis of the intervallic rivalry model is based on the different levels of rarity or ubiquity of intervals within specific intervallic categories (interval classes) in the diatonic set—that is, a representation of all possible intervals among the seven tones of any major scale, independent of ordering or spelling. Figure 1 provides background for understanding the rare-interval hypothesis. In Figure 1a–1c, all possible intervals between pairs of pitches within the C major diatonic set are shown in musical notation. These intervallic relationships are characteristic of all major diatonic sets and hence function as invariant properties that apply to all transpositions of the C major set. The 21 intervals of the diatonic set are represented in six different sizes; these, in turn, are grouped into three categories based on pairs of intervals with their common musical labels, seconds (one or two semitones), thirds (three or four semitones), and fourths (five or six semitones).\textsuperscript{1} The two sizes of intervals within each category are notated such that the less frequent interval in each pair is shown with a filled notehead and the more frequent with an open notehead. For example, in the “seconds” category, Figure 1a shows the two intervals of the diatonic set with a one-semitone span (minor seconds) with filled noteheads and the five intervals with a two-semitone span (major seconds) by open noteheads. Similarly, in the “thirds” category in Figure 1b, there are four three-semitone spans (minor thirds) and three four-semitone spans (major thirds). In the “fourths” category in Figure 1c, there are six five-semitone spans (perfect fourths), and a single span of six semitones (augmented fourth).\textsuperscript{2} Notice that in the “fourths” category, not only is the augmented fourth the less frequently occurring interval of this categorical pair, it represents the least frequently occurring interval class within the whole diatonic set of 21 intervals. The largest and rarest interval class in the set spans one-half octave and thus inverts into another half-octave. When inverted, the augmented fourth is the smallest fifth in the set (diminished); when neither spelling is specified, it is referred to as a tritone. Figure 1d portrays a relative frequency distribution of interval size in which each of the six interval classes in the diatonic set is

1. Within interval classes, it is not intervals indicated but numbers of semitones. Correspondingly in Figure 1: (1) Unisons and octaves are not represented because these intervals possess the same tone chroma; thus, they do not provide any information about the pitch set that a single tone has not already conveyed. (2) Intervals larger than the augmented fourth and up to the octave (fifths, sixths, and sevenths) are not represented because they may be generated by inversion—moving the lower note up, or the upper note down, by one octave in each interval. Because inversion provides only redundant information, each interval larger than one-half octave is represented by the number of semitones its inversion comprises. (3) Any enharmonic spelling of an interval is represented by the number of semitones it spans.

2. It should be noted that the rarity or ubiquity of interval classes in the diatonic set is not necessarily related to the frequency or infrequency with which intervals in those classes are found in tonal musical compositions.
associated with a unique relative frequency score. This diatonic set property has been referred to as unique multiplicity (Babbitt, 1960). It refers to the fact that each interval conveys a distinct relative number of occurrences. In this context, those intervals with a relative frequency of 0.10 or less are defined as rare intervals (see the shaded area of Figure 1d). Thus two intervals qualify as rare intervals: the two minor seconds and the single tritone.

Gauldin (1983) has suggested that the canonical diatonic set developed as a basis of tonality because of the aural implications of the property of unique multiplicity. The rare-interval hypothesis builds on the idea that common and rare intervals afford different kinds of tonal information to the listener. It assumes that a listener's identification of important tonal relationships depends on a (usually tacit) knowledge that the six interval classes are not only uniquely paired with certain frequencies of occurrence in a diatonic set but that these relative frequencies also signal the reliability with which a given interval can be located within the set. That is, a major third is always built on scale degrees 1, 4, and 5 in a diatonic set. The more common intervals (major seconds and perfect fourths) correspond to more locations in the set and as a consequence are difficult to place. The rarer intervals (minor seconds and the tritone) are easier to place. For example, the tritone occurs only between degrees 4 and 7 in the major set. Browne (1981) uses the term position finding to describe listening behavior that is sensitive to these kinds of degree-related aural implications of intervals. He argues that position finding contributes to a listener's identification of the tonal function of the pitches involved. The rare-interval hypothesis reflects the assumption that rare intervals facilitate position finding and thus key discovery. Mapping of the pitches in a musical composition onto scale degrees in a diatonic set is fraught with minimal ambiguity when these pitches outline a rare interval within the set. This can be illustrated with each of the two rare intervals shown in Figure 1. The minor second occurs only twice in a diatonic set (between degrees 3–4 and 7–8). Therefore, when a musical event embeds a third tone related by a minor second to tones forming more common intervals, this combination facilitates a listener's identification of the tonal functions of all three pitches because it can be shown that only two key interpretations are possible (within the major mode and barring chromaticism). The tritone occurs only once in a set, but because it inverts into another tritone, it also occurs in an enharmonic respelling in another key; that other key itself is a tritone from the first key. But finally, a musical event that adds a third tone to a tritone renders position ambiguity minimal, facilitating key identification, disambiguating the spelling of the tritone, and thus uniquely specifying a major-mode tonal center.

Figure 2 illustrates the basis on which the rare-interval hypothesis rests.
The C major set is produced by gray and white keys. The tritone lies between F (or E♯) and B, the gray-shaded keys. The black keys, when combined with the two gray pitches of the tritone, also form a diatonic set, the F♯ major set (F♯-G♯-A♯-B-C♯-D♯-E♯). These two diatonic sets are mutually exclusive except for the tritone. When the tritone is added to any third tone, this determines whether the key is C major or F♯ major. While the more commonly occurring intervals serve other tonal functions in the diatonic set, they do not specify major key. The intervallic rivalry model includes the recognition that listeners tacitly learn the disambiguating potential of rare intervals and come to rely on this sort of information to identify key. Support for the rare-interval hypothesis is found in Brown and Butler's (1981) study, which demonstrated that trained listeners can reliably determine tonal centers for the major mode after hearing only three tones, two of which span a tritone.

The temporal-order hypothesis includes the distinctive implications that rare intervals present as a function of the order in which the pitches forming these intervals occur. This hypothesis is proposed to account for differences a listener would experience as a function of hearing an unordered collection of tones (a pitch set) versus a melody, as, for example, in the Christmas carol, O Tannenbaum. In the melody, where a clear sense of key obtains, the order of pitches of the tritone occurs with the fourth scale degree followed by the seventh, an ordering that should effectively imply motion toward tonic for the listener because it implies an arpeggiation of dominant-seventh chord resolving to tonic (V7 → I). This sets up an expectancy for goal-directed harmonic motion similar to that in other typical musical contexts (Jones, 1976, 1981). When the temporal ordering of a rare interval does not suggest such specific motions, the listener's sense of goal completion and, hence, key discovery, is weaker. Support for this hypothesis is found in studies showing that listeners' judgments of tonal centers were clearer in patterns in which the subdominant preceded the leading tone (scale degree 4–7) than the reverse (Boltz, 1989; Brown,
1988; Brown & Butler, 1981; Butler, 1983). These same studies have also illuminated ways in which listeners make tonal assumptions from incomplete tonal evidence.

In summary, the intervalllic rivalry model is an explanation of how listeners are able to use information presented in the two dimensions of music—pitch and time—in order to track the path of tonality in music: from initial establishment of a tonal center through the network of modulations and briefer tonal excursions to the overall sense of the key scheme of a work.

Others have focused on the importance of the more common intervals within the diatonic set as a source of reinforcing an established key. Krumhansl's tonal hierarchy model describes the hierarchic nature of perceived similarity and stability of each tone from the chromatic set in tonal contexts in terms that emphasize the salience of certain pitches, those that occur in intervals of thirds and fifths (Krumhansl, 1979, 1990a; Krumhansl & Shepard, 1979). Much evidence, derived from pairwise similarity judgments of single tones within a tonal context, suggests that the tonal hierarchy can be represented as a conical configuration with the tonal center, considered to be the most stable of tones in a key, located at the vertex of the cone (Krumhansl, 1979). By virtue of their interpretations as tones of the tonic triad, the mediant and the dominant form a closely related cluster near the vertex. Remaining diatonic tones form a less closely related subset farther from the vertex, and finally chromatic tones are farthest from the vertex and regarded as least stable. Thus, in a C major tonal context, the pitches C, E, and G are assumed most stable in this hierarchy, whereas those outside the diatonic set are least stable (Krumhansl, 1979; see also Krumhansl, 1990a, p. 128). Other support for this view comes from “good completion” ratings of single tones (probes) that followed a specified tonal context (e.g., a scale). For example, Krumhansl and Shepard (1979) found that when the probe tone was the tonic, listeners rated it the best-fitting completion of all tones heard after a scalar pattern context. Other pitches of the tonic triad were rated next best as completions and remaining diatonic pitches were judged to fit better than nondiatonic ones.

The tonal hierarchy has been expanded into a harmonic hierarchy describing tonal relationships among chords (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha, & Kessler, 1982) and among keys (Krumhansl & Kessler, 1982). It is also supported by strong correlations of hierarchical strength with frequency of occurrence of tones within tonal musical compositions (Krumhansl, 1987, 1990a) and has formed the basis of research into related problems as well (e.g., Cuddy & Badertscher, 1987; Lerdahl, 1988; Sundberg & Frydén, 1987). The tonal hierarchy has been influential as a psychological model of pitch space and is widely considered to be a powerful cognitive map of a listener’s tonal pitch relations.
In both the intervocal rivalry model and the tonal hierarchy model, ordinary listeners are able to make tonal decisions quickly from partial and fleeting musical evidence. A crucial difference between the two approaches involves the role of tonal context. In the tonal hierarchy model, it is assumed that incoming contextual information is sufficient to specify key for subjects who then render similarity, completion, or "good fit" ratings. Thus, the hierarchy reflects an enculturated listener's application of musical knowledge based on a given specified tonal context. It emphasizes the stability of certain prominent pitches—tonic, mediant, dominant—that happen to be outlined by those interval classes more common in the diatonic set. The strength of the tonal hierarchy model lies in descriptions of ways these common intervals reinforce key. By reinforcing key, we mean: Given a fixed cognitive anchor pitch, the presentation of more-common intervals of the diatonic set (major seconds, minor and major thirds, and perfect fourths) supports tonal functions within the tonal framework of the anchor pitch. Their reinforcing functions contribute importantly to the jagged tonal profiles reported by Krumhansl and her colleagues.

By contrast, in the intervocal rivalry model, it is assumed that enculturated listeners respond to incoming context by engaging in an active discovery process oriented toward identifying an initial tonal center and identifying evidence for change of tonal center over the course of music involving modulations. In this process, rivals for tonal center are successively eliminated in real time as a musical work unfolds. The emphasis of this model complements that of the tonal hierarchy in that it encompases the roles played by less common intervals. Rare intervals and their orderings are assumed to operate initially over time to contribute to tonal profiles by fixing a single pitch as a tonal center. Therefore, the strength of the intervocal rivalry model lies in suggesting how rare intervals facilitate discovery of key.

The two models address different aspects of tacit knowledge about tonality. Here it is assumed that key discovery precedes tonal reinforcement but that both processes are necessary for a listener to follow tonal music in real time. This paper focuses on the issue of key discovery in experiments overtly requiring completion judgments from listeners.

Figure 3 serves to illustrate aspects of the discovery process implied by the intervocal rivalry model. According to the primacy hypothesis, a listener encountering the pattern in Figure 3a in real time is initially biased to consider F, the first tone heard, as the tonic. This hypothesis cannot be confirmed without additional musical evidence. The major third between F and A is the least frequent interval of the diatonic set appearing in the pattern. The smallest number of possible major keys specified by the major third is three, meaning that F and A are diatonic in three different major keys (shown in the frequency distribution in Figure 1d): C major, F
major, and B major. All three are rivals for tonic from the perspective of the listener. The major triad cannot unambiguously confirm F as the major tonal center; either C or Bb can replace F as a prime candidate. The note F is a plausible tonic until additional pitch information arrives to persuade the listener that there is an alternative. Additional evidence can also keep the sense of F as tonal center alive. For example, Figure 3b shows E as a fourth tone; this introduces a rare interval, the minor second (indicated by a bracket), narrowing the feasible candidates for a major tonal center to F and C. However, if the fourth tone were B and not E, this would shift the balance in favor of C major because the interval formed by B and F is a tritone (bracketed in Figure 3c). The single additional tone necessary to provide the minimal intervalllic context needed to prevent enharmonic misinterpretation of the tritone could, in this example, be either the A or the C; it need not be the tonic. In short, a listener is viewed as evaluating rival contenders for tonic as a musical event unfolds; the primacy hypothesis implies a bias for the first tone, but as long as the sequence "fits" in other keys, other logical candidates can also be entertained. Eventually, rare intervals permit the listener to determine the appropriate keynote, tacitly discarding alternatives refuted by unfolding evidence.

The example above illustrates the winnowing of rival hypotheses in the case where the rare intervals (Figure 3b and 3c) occur in optimal temporal orders. Typically, the tritone is found in harmonic groupings that have a dominant function: dominant seventh chord and the diminished triad built on the leading tone. Both chords are considered strong signals of an impending tonic harmony. Each represents the harmonic function of dominant, that which most strongly influences the listener's assessment of tonal center. Of course, the dominant tone as the fifth scale degree also participates in the tonic triad, meaning that it can also be interpreted as contributing to a sense of completion or stability (e.g., as in the tonal hierarchy). But traditionally the dominant has also been taken to represent the polar opposite of the stability of tonic function. Both of these interpretations are captured in the role attributed to the dominant scale degree in the intervalllic rivalry model. The function of dominant, as contrasting with and,
indeed, setting up motion leading to, the tonic has been a basic tenet of Western music theory for nearly three centuries (Rameau, 1722/1971).

The contrasting functions of the dominant, assumed to underlie an implied sense of continuation of harmonic motion from dominant to tonic harmony, support the temporal-order hypothesis of the intervocal rivalry model. Thus, for example, if the tones of Figure 3c were reordered as C-B-A-F, as in Figure 3d, the likelihood that the listener would finally choose C as the tonic would decrease, in spite of the compelling contour opening with C. Primarily, this is because the tritone is presented in the reverse of its optimal ordering for the key of C.

The temporal-order hypothesis can be used to examine differences in good completion ratings as a function of orderings of pitch events associated with rare intervals. This can be illustrated in Figure 4, which shows different temporal orderings of a tritone as found in the diminished triad. If tones comprising a diminished triad (e.g., B-D-F) were presented in a temporal arrangement mimicking events commonly encountered in tonal melodies, one might expect that a listener would be led to rate more highly those orderings that embody a dominant harmonic motion strongly implying the leading-tone-to-tonic resolution (B to C in C major). If the tones of the triad were reversed, this ordering might be less likely to provoke the listener to identify a subsequent probe tone of C as a good completion. Instead, F# may be favored. Each interpretation is a voicing of an implied dominant-tonic progression—but in a different key. The first four tones in the upper staff in both Figure 4a and 4b would sound identical in isolation. In Figure 4a, however, the harmonic accompaniment allows the bracketed tones B and F to form the temporal succession of scale degree 7 to scale degree 4; this tritone resolves typically—F as seventh of the chord moves down by step to the final melodic tone E, the third of the tonic triad, as the leading tone B, in the implied inner voice, resolves upward by halfstep to the tonic. In Figure 4b, the diminished triad has been reharmonized in the key of F# minor. The four tones (indicated by a square

![Figure 4](image_url)

*Fig. 4.* One tonal interpretation of the arpeggiated diminished triad, harmonized as the dominant-seventh chord in the key of C major. (b) Another interpretation suggests an enharmonic respelling of the same tritone, harmonized as the dominant seventh chord in F# minor. Numbers with carets above them indicate scale degrees.
are diatonic in this key where E# —now the leading tone, also resolves typically with B now functioning as the seventh of the chord, resolving downward by step. In other words, the members of the tritone interval have switched roles in this example: the subdominant from Figure 4a is now the leading tone, and vice versa. The question is how listeners might evaluate completion in ambiguous melodic arpeggiations in the absence of overt harmonic reinforcement.

The goal of the present experiments is to clarify respective roles of common and rare intervals in listeners’ completion judgments for individual tones in certain tonal contexts and particular temporal orderings. It has been suggested that rare intervals, even in musical contexts, are not sufficient to establish key for listeners, at least ordinary listeners (e.g., Cuddy & Badertscher, 1987; Krumhansl, 1990b). Their evidence derives primarily from probe-tone tasks where completion judgments about 12 probe pitches (the chromatic pitch classes) presented in three different melodic contexts are compared: a major scale, a major triad, and a diminished triad. Cuddy and Badertscher reported profiles of probe-tone ratings (called probe-tone profiles) as negative evidence for the intervallic rivalry model: Their major triadic context appeared to yield profiles consistent with establishment of tonality, whereas this was not the case with the diminished triadic pattern. Specifically, jagged probe-tone profiles emerged for the major triadic context in which probe tones that were diatonic pitches, especially those interpreted scale degrees 1, 3, 4, and 5, were given high ratings. Performance with the complete major scale yielded a jagged profile comparable to that of the major triad, except that ratings for the tonic probe were much higher and those for the subdominant relatively lower. Finally, and most relevant for the intervallic rivalry model, were their findings based on the second triadic context, the diminished triad. Here the jagged tone profile essentially disappeared. No single probe, including the major-mode tonic, was rated distinctively higher than others, a pattern the authors interpreted to mean that the rare intervals embedded within this pattern did not allow listeners to “recover” a tonal hierarchy (Cuddy & Badertscher, 1987, p. 618).

At least two aspects of Cuddy and Badertscher’s approach might account for their findings with respect to the diminished-triad context. Both relate to the underlying rationale of the intervallic rivalry model, which is based on the idea of implied harmonic motion over time. The first concerns the reliance on jagged profiles of completion ratings as determinate evidence for key identification. Butler (1989) has cautioned that use of completion ratings as unambiguous or conclusive evidence for key identification is risky. But if we assume that strong completion judgments can correspond to tonic resolution, among other things, then they provide partial evidence for inferences about tonic. With these caveats, the intervallic rivalry model makes certain predictions about completion ratings: If
a listener is presented with a diminished triad and if it sets up a specific goal that can be satisfied only by a single pitch, namely the correct tonic, then only that pitch should receive a very high rating score in a completion task. In particular, other pitches within the dominant harmony (e.g., G in C major) should be heard as continuing the dominant harmony and so may indeed seem to be weak completions. Thus, because the intervallic rivalry model emphasizes information that influences initial discovery, it is not aptly validated by the same jagged tonal profile that the tonal hierarchy predicts. In principle, it does not predict the “recovery” of the tonal hierarchy, as such, in all contexts, and this is because intervallic rivalry model predictions about completion judgments depend on the degree of ambiguity in the presented context. According to this model, contexts that embed more rival candidates for the tonic will produce more peaks in a probe-tone profile because one important determinant of “good completion” for a probe is that it function as a possible tonal center. The second aspect of the Cuddy and Badertscher investigation that might have accounted for their failure to find strong completion ratings for the tonic probe within the diminished-triadic context involves the ordering of pitches within the triadic pattern they used. In their triadic pattern, the pitches outlining the tritone appeared in reversed order, meaning that it was as likely to imply another key as the correct one. Together these two features of their studies can account for the apparently poor performance of the intervallic rivalry model. From the perspective of the intervallic rivalry model, the findings Cuddy and Badertscher reported are not entirely surprising. However, they raise the question of whether or not listeners do, in fact, rate the correct keynote as rendering a very good completion to a diminished triad when circumstances are more favorable for this, given the assumptions of the intervallic rivalry model. The present report addresses this question.

Two experiments comprise the present study. Experiment 1 was designed to replicate the findings of Cuddy and Badertscher; Experiment 2, to explore more pointed predictions of the intervallic rivalry model associated with the temporal-order hypothesis. In both experiments, listeners with different degrees of musical training were presented with three different melodic patterns: an arpeggiated major triad, an ascending major scale, and an arpeggiated diminished triad. They were asked to judge how well a probe tone from the chromatic set completed each pattern. In Experiment 1, these context patterns were presented in the temporal ordering chosen by Cuddy and Badertscher, an ordering predicted by the intervallic rivalry model to be optimal for the major triadic and scalar contexts for discovery of the same major key assumed by the tonal hierarchy model; an ordering least favorable for the diminished triad context by virtue of its last tone, the subdominant. Experiment 2 presented each of
the three patterns in a different temporal order, an ordering predicted by
the intervallic rivalry model to be optimal for specific keys for both triadic
contexts (F major for the major triad and C major for the diminished
triad) and to be neutral for the scalar context. It was expected that the
pattern of findings of Cuddy and Badertscher would be replicated in Ex-
periment 1, but not in Experiment 2.

Experiment 1

Listeners in this experiment rated probe tones from the chromatic set in
the context of three different melodic patterns, shown in Figure 5. These
patterns are identical to those used by Cuddy and Badertscher (1987). 
Figure 5a is an arpeggiated major triadic pattern, which embeds two of the
potential keynotes in major mode, C repeated as first and third notes and
G as last tone. Figure 5b is an ascending major scalar pattern, which
repeats the single keynote C as both first and last tones, an octave apart.
Figure 5c is an arpeggiated diminished triadic pattern with a contour
delete identical to that of the major triadic pattern. Unlike the patterns in Figure
5a and 5b, the pattern in Figure 5c contains no major-mode keynote.

The intervallic rivalry model makes several predictions relating comple-
tion judgments and the shape of probe-tone profiles in these contexts.
First, with respect to the scalar pattern, all three hypotheses—
primacy, rare-interval, temporal-order—converge to predict that C will
emerge as the strongest candidate for good completion and tonal center.
The clear presentation of both the tones in the context and the common
intervals they outline should reinforce the key and provide support for a
jagged profile consistent with predictions of the tonal hierarchy model.

Theoretically, the more interesting predictions involve the two triadic
patterns. According to the intervallic rivalry model, the common intervals
that comprise a major triad render it ambiguous tonally, and therefore it
cannot uniquely specify a keynote. Nevertheless, by the primary hypo-
thesis, the listener should be biased to treat C, the first tone heard in this
presentation, as a keynote. Assuming that any candidate for a tonic will
receive a strong completion rating, in contexts where several candidates
abound, the intervallic rivalry model grants a basis for predicting a jagged

Fig. 5. Stimuli used by Cuddy and Badertscher (1987) and in Experiment 1 of the present
study. The probe tone was heard at the point corresponding to the rectangle in each
example, a–c.
profile. In the case of the major triad, it predicts that, to the extent that ratings of rival tonic candidates contribute to the jaggedness of a resulting profile, relatively higher profile peaks will occur for those probe tones.

Finally, of special importance are predictions of the intervallic rivalry model for probes following the diminished triadic context, for which a different probe-tone profile is predicted as evidence for key discovery. Here the mere presence of the rare interval (between B and F) with the additional tone (D) would incline listeners to rate only C as a good completion in major mode. However, the temporal-order hypothesis predicts that the diminished triadic context provides clearer information about key when the ordering of pitches that outline the tritone implies voice-leading typical in a dominant-to-tonic progression. If the temporal order is not directed toward tonic (as it is not in Experiment 1, the pattern ending with the subdominant), then listeners’ judgments of best completion might not coincide with their sense of key for the pattern. When tonal relationships between the tones of the pattern as ordered and the probe tone are evaluated, listeners’ good completion ratings could be influenced by the voice-leading considerations discussed earlier: implication of a scale-degree 4 to 3 resolution of the tritone in an upper voice (the C major tonal interpretation notated in Figure 4a) or implication of a leading-tone to tonic resolution (the F# minor interpretation in Figure 4b). That is, the tritone can be interpreted as either of the particular voicings of the implied dominant-to-tonic progressions in Figure 4. In Experiment 1, this means that probe tones rated higher than others should be C, E, and F#.

Method

SUBJECTS

The subjects were 45 adults with various levels of musical experience. The study group included 13 women and 32 men from 18 to 43 years old, with a mean age of 24.5 years (median, 22 years). Cuddy and Badertscher divided their subjects into groups according to their grade level (from no grade through Grade IX) in the Royal Conservatory examination, but this system does not exist in the United States. Therefore, subjects in Experiment 1 were divided into three groups on the basis of their performance on a test of their ability to recognize tonal centers from excerpts from tonal music. The high group had a mean number of 8.1 years of private lessons on an instrument or voice; the medium group, 6.8 years; and the low group, 5.9 years.

APPARATUS

Stimuli were produced by a MicroVax II computer using Cmusic (© 1988, Regents of the University of California) routines for tone generation and sequencing. The sampling rate was 44.1 kHz. Each stimulus tone comprised nine octave-related sinusoidal components within the range of 15–8800 Hz. These tones, commonly called “Shepard tones” (after Shepard, 1964), may be described as a raised inverted cosine envelope over a logarithmic frequency abscissa. They have the practical advantage of diminishing octave effects
between melodic intervals. Signals were recorded on a Sony SL-HF 450 tape transport controlled by a Sony PCM 601 esD digital controller, then copied through a Nakamichi 480 audiocassette recorder onto 3M AVX 46 tape. Stimuli were presented binaurally through AKG K 240 headphones. Subjects adjusted volume to a comfortable level.

STIMULI

Stimuli were faithful replications of those used by Cuddy and Badertscher (1987). Three types of melodic patterns—a major triad, a major scale, and a diminished triad—followed by probe tones that were a random presentation of each member of the 12-tone equal-tempered (A_{440}) pitch set, comprised the 72 stimuli. The patterns were extracted from three different diatonic sets: B, C, and C# major. These patterns are shown in Figure 5; to facilitate direct comparison, all figures and tables in this article are shown transposed to C major. Pitch class numbers 0–11 correspond to an ascending chromatic scale where C = 0, C# = 1, . . . , B = 11. Each tone in the patterns had a duration of 0.33 s, each probe tone had a duration of 1.00 s, and both had a rise and fall time of 20 ms. Probe tones were separated from patterns by a 0.1-s gap. For counterbalance, there were six blocks of stimuli, two of each of the three types, containing different random orderings of the 12 probe tones.

DESIGN

The design was a 3 (Musical Experience) × 3 (Pattern) × 12 (Probe Tone) mixed factorial. The three levels of subjects’ musical experience (high, medium, low) formed a between-groups factor; within-subject factors were three context patterns (major triad, major scale, diminished triad) and 12 probe tones (pitch classes 0–11).

PROCEDURE

Each subject heard six blocks of trials, two for each of the three types of patterns. The order of patterns was randomized across subjects, who were tested individually. The subjects’ task was to rate how well each probe tone “provided a musical completion” for the preceding pattern. The rating “1” signified “sounds good”; “7” signified “sounds bad.” Each subject filled out a background questionnaire, participated in the pretest, and heard five practice trials before beginning the experiment. Each subject read the following instructions:

In this music perception experiment, the experimenters would like to have your honest, subjective evaluations of how well the FINAL note of each pattern provides a MUSICAL COMPLETION to the pattern. Your evaluations are to be circled on the 7-point rating scales provided below, with 1 meaning sounds good and 7 meaning sounds bad. You will hear and rate 5 practice trials, followed by 72 experimental trials.

Each session lasted approximately 25 min.

Results

OVERALL ANALYSIS

Probe-tone judgments are reported here as profiles (Krumhansl & Kessler, 1982) for each pattern, transposed to the key of C major. The
arpeggiated major triadic pattern, which can function as the triad built on the first, fourth, or fifth scale degree in major keys, is shown in its interpretation as the first—the tonic triad in C major. The scale is the C major scale. The only diminished triad diatonic to a major key is built on the leading tone, the seventh scale step; therefore, the diminished triadic pattern with the root B is also shown as interpreted in a C-major profile.

The averaged probe-tone ratings for each pattern for each musical experience group are shown in Table 1. The groups who differed in training and pretest performance did not differ significantly overall \([F(1, 42) = 0.78]\). They did not respond differently to the three patterns, and the interaction of pattern with group was insignificant \([F(4, 84) = 1.09]\). A three-way interaction of group with pattern and probe tone also did not prove significant \([F(44, 924) = 0.88]\).

Profiles (collapsed over musical experience groups) for the three patterns are shown in Figure 6, in which 1 is the highest rating and 7 is the lowest rating. There was a main effect of pattern \([F(2, 84) = 13.42, \ p < .001]\) with mean ratings for the diminished triad \((M = 3.71)\) higher than for the major triad \((M = 3.81)\), and for the major triad higher than for the major scale \((M = 4.22)\). Probe tone also had a significant overall influence on ratings \([F(11, 462) = 78.35, \ p < .0001]\), the nature of which is evident from the jagged profiles in Figure 6. Pattern exerted a differential effect on probe-tone profiles, leading to a significant interaction of these two variables \([F(22, 924) = 9.46, \ p < .00001]\). Details of this interaction are considered in separate profile analyses of the three patterns.

### TABLE 1

<table>
<thead>
<tr>
<th>Probe Tone</th>
<th>C (0)</th>
<th>C♯ (1)</th>
<th>D (2)</th>
<th>D♯ (3)</th>
<th>E (4)</th>
<th>F (5)</th>
<th>F♯ (6)</th>
<th>G (7)</th>
<th>G♯ (8)</th>
<th>A (9)</th>
<th>A♯ (10)</th>
<th>B (11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.93</td>
<td>5.03</td>
<td>4.03</td>
<td>4.50</td>
<td>3.20</td>
<td>3.07</td>
<td>4.40</td>
<td>2.60</td>
<td>3.80</td>
<td>3.33</td>
<td>4.00</td>
<td>4.33</td>
</tr>
<tr>
<td>Medium</td>
<td>1.83</td>
<td>5.50</td>
<td>4.03</td>
<td>5.00</td>
<td>3.23</td>
<td>3.13</td>
<td>5.20</td>
<td>2.10</td>
<td>4.37</td>
<td>3.70</td>
<td>4.20</td>
<td>4.90</td>
</tr>
<tr>
<td>High</td>
<td>1.23</td>
<td>5.70</td>
<td>3.80</td>
<td>5.00</td>
<td>2.93</td>
<td>3.23</td>
<td>4.87</td>
<td>2.03</td>
<td>4.33</td>
<td>3.37</td>
<td>4.00</td>
<td>5.13</td>
</tr>
<tr>
<td><strong>Major Triad</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.67</td>
<td>4.83</td>
<td>3.73</td>
<td>4.90</td>
<td>3.73</td>
<td>4.03</td>
<td>5.30</td>
<td>3.30</td>
<td>5.20</td>
<td>3.57</td>
<td>5.13</td>
<td>3.57</td>
</tr>
<tr>
<td>Medium</td>
<td>1.30</td>
<td>5.60</td>
<td>4.13</td>
<td>5.47</td>
<td>4.40</td>
<td>4.17</td>
<td>5.40</td>
<td>3.47</td>
<td>5.83</td>
<td>4.23</td>
<td>5.23</td>
<td>4.10</td>
</tr>
<tr>
<td>High</td>
<td>1.43</td>
<td>5.87</td>
<td>3.63</td>
<td>5.53</td>
<td>3.53</td>
<td>4.10</td>
<td>5.07</td>
<td>2.63</td>
<td>5.03</td>
<td>3.60</td>
<td>5.23</td>
<td>3.93</td>
</tr>
<tr>
<td><strong>Diminished Triad</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2.13</td>
<td>4.37</td>
<td>3.40</td>
<td>3.63</td>
<td>3.10</td>
<td>2.87</td>
<td>3.77</td>
<td>3.20</td>
<td>4.23</td>
<td>3.60</td>
<td>3.50</td>
<td>3.30</td>
</tr>
<tr>
<td>Medium</td>
<td>2.17</td>
<td>4.90</td>
<td>3.90</td>
<td>4.47</td>
<td>3.67</td>
<td>3.23</td>
<td>4.03</td>
<td>3.10</td>
<td>4.17</td>
<td>4.27</td>
<td>4.27</td>
<td>3.40</td>
</tr>
<tr>
<td>High</td>
<td>1.83</td>
<td>5.40</td>
<td>3.70</td>
<td>5.03</td>
<td>3.33</td>
<td>3.57</td>
<td>3.93</td>
<td>2.53</td>
<td>5.03</td>
<td>4.87</td>
<td>4.67</td>
<td>3.03</td>
</tr>
</tbody>
</table>
PROFILE ANALYSES

Separate analyses of variance were performed on the profiles for each of the three pattern contexts shown in Figure 6. Figure 7 shows corresponding profiles from Cuddy and Badertscher’s experiment.

Major Triadic Pattern

The profile of responses observed for the arpeggiated major triadic pattern was consistent with findings reported by Cuddy and Badertscher (Figure 7a), yielding significance in these comparisons: C rated higher than G \( [F(1,42) = 9.54, p < .005] \) and together with G, these two probes (assumed tonic plus dominant) were judged as better completions than the median E \( [F(1,42) = 41.90, p < .0001] \). Overall these three tones of the major triad, the only tones in the presented context pattern (C, E, G—pitch classes 0, 4, 7) were rated higher than the other tones in the diatonic set of C major (D, F, A, B—pitch classes 2, 5, 9, 11) \( [F(1,42) = 9.68, p < .005] \). Of the latter, the probe tone F (pitch class 5) stands out as different from other diatonic tones in that it was judged to be a significantly better completion than the remaining diatonic tones, none of which appeared in the context pattern \( [F(1,42) = 23.79, p < .0001] \). In fact, the mean completion ratings for F (3.14) were comparable to those for E (3.12). Furthermore, E was rated as a significantly better completion than other nontriadic diatonic tones (D, A, B averaged) \( [F(1,42) = 24.80, p < .0001] \). Of the nontriadic diatonic tones, the leading tone, B (pitch class 11) was

![Graph](image)

Fig. 6. Comparison of profiles from Experiment 1 (Replication): C major triadic pattern, C major scalar pattern, and B diminished triadic pattern. pc = pitch class.
Fig. 7. C major profiles from Experiment 1 (replication), shown overlaid with profiles from Cuddy and Badertscher (1987). Pitch classes 0–11 = C–B; on the rating scale, 1 = "sounds good" and 7 = "sounds bad." (a) Profile in response to the C major triadic pattern; (b) profile in response to the C major scalar pattern; and (c) profile in response to the B diminished triadic pattern.
judged the poorest completion (4.79), significantly poorer than the diatonic probe with the next highest rating, namely A (pitch class 9) \(F(1,42) = 43.29, p < .0001\). Finally, as a set, the seven diatonic pitches were rated as significantly better completions than the remaining five chromatic ones in the 12-tone set \(F(1,42) = 26.56, p < .0001\).

These results amount to a complete replication of Cuddy and Badertscher’s results for this pattern. Correlations confirm the similarity of probe profiles across the two experiments (Figure 7; \(r = .960, p < .0001\)).

**Major Scalar Pattern**

The profile obtained from responses to the ascending major scalar pattern resembles that found by Cuddy and Badertscher in several important respects (see Figure 7b). The tonic C was rated higher than the dominant G \(F(1,42) = 59.34, p < .0001\) and, when taken together, tonic and dominant were again rated higher than the mediant E \(F(1,42) = 92.77, p < .0001\). Tones of the tonic triad, taken together, were judged to be significantly better completions than other diatonic tones \(F(1,42) = 5.38, p < .05\). In contrast to the major triadic context, in this context, neither F nor E (considered separately) was judged as a significantly better completion than nontonic-triad diatonic tones (D, A, B). Mean completion rating for the latter group was 3.83; in this context, both E and F received slightly higher scores (i.e., 3.89 and 4.07, respectively), which suggests that these pitches were heard as slightly poorer completions. Finally, of all of the tones presented within the scalar pattern, the diatonic pitches were rated higher than chromatic notes, which were not heard \(F(1,42) = 47.31, p < .0001\).

Some comparisons that Cuddy and Badertscher found to be significant, however, did not emerge as such in the present study. The most notable difference concerned F as probe tone. Their listeners judged F to be a significantly poorer completion than other nontonic-triad diatonic tones in this context. However, in the present study, the probe tone F yielded completion ratings only slightly poorer than those found when the three nontonic-triad diatonic probe tones (D, A, B) were averaged. This difference was not significant \(F(1,42) = 2.36, p > .10\). Furthermore, neither the pairwise comparisons of A versus B nor D# versus C# was significant \(p > .10\).

Again, overall the comparisons with the C major profile recovered by Cuddy and Badertscher confirm the general similarity among the profiles found in the two experiments \(r = .940, p < .0001\).

**Diminished Triadic Pattern**

The profile for the diminished triadic pattern in this experiment exhibits some important differences from that reported by Cuddy and
Badertscher (Figure 7c). The most importance difference, theoretically, is the finding that the major-mode tonic C is rated as a substantially good completion in this experiment with a rating that differs relatively more from those of other probes than evident in profiles of Cuddy and Badertscher. C was rated higher than the dominant G \( [F(1,42) = 13.58, p < .0001] \) and together tonic and dominant were judged better completions than the median E \( [F(1,42) = 22.42, p < .0001] \). Another difference from the Cuddy and Badertscher finding involved the F probe, which was rated a better completion than other diatonic tones not contained in the tonic triad \( [F(1,42) = 9.32, p < .004] \). All diatonic tones were rated higher than chromatic ones \( [F(1,42) = 12.41, p < .005] \) and within the diatonic set, B was rated significantly higher than A \( [F(1,42) = 15.18, p < .0003] \). Other differences involved chromatic tones: F# was rated significantly higher than the average of C#, D#, and G# \( [F(1,42) = 12.35, p < .005] \); and C# was rated higher than D# \( [F(1,42) = 4.23, p < .05] \), but D# together with F# and G# were rated higher than C# \( [F(1,42) = 16.89, p < .0002] \).

These detailed findings as well as a correlational analysis of the profiles in Figure 7c confirm that, overall, performance with the diminished triadic pattern differs from that reported by Cuddy and Badertscher. The correlation with profiles of Cuddy and Badertscher for the diminished triad was .437 and this was not statistically significant \( (p < .156) \). On the other hand, the probe-tone profile of Figure 7c for Experiment 1 does correlate fairly well with that resulting from the C major triadic pattern of Experiment 1 \( (r = .796, p < .001) \). The latter finding is clearly the result of generally higher ratings for diatonic pitches in both triadic contexts as well as higher ratings for C, E, F, and G.

**Discussion**

Overall, there is substantial comparability between the results of Experiment 1 and those reported by Cuddy and Badertscher (1987), as seen in Figure 7. Completion judgments of listeners of all three training levels are striking in their similarities across these two experiments, particularly for the major scalar pattern (Table 1, Figure 7). The jagged tone profiles found in this context are consistent with those reported not only by Cuddy and Badertscher but by others (Krumhansl, 1990a; Krumhansl & Shepard, 1979). Most listeners rated the (assumed) tonic as the best completion and preferred probes that belong to the diatonic set over those considered chromatic as completions. The finding that major scalar patterns produce evidence consistent with both discovery of the tonal center and sensitivity to major tonal functions within the established and related keys not only
replicates Cuddy and Badertscher’s study but is also generally consistent with Krumhansl’s predictions based on a tonal hierarchy analysis (Krumhansl, 1990a). One departure from tonal hierarchy predictions in this context concerns judgments of the mediant pitch (E), which listeners rated to be roughly as good a completion as other nontonic-triad diatonic pitches sounded within the scalar pattern (e.g., A, B). According to this view, judgments of E should reflect significantly stronger (lower) ratings than those found with any nontonic-triad diatonic members; it did not.

With respect to profiles associated with the two triadic contexts, although the present data and those of Cuddy and Badertscher are in substantial agreement, the disparities are interesting in light of different theoretical interpretations of the shape of probe-tone profiles. The intervallic rivalry model and the tonal hierarchy model predict different patterns of peaks for the two triadic contexts.

Consider first the major triadic pattern. In this context, the tonal hierarchy model predicts that scale degrees 1, 3, and 5 should be rated as most stable, whereas the intervallic rivalry model predicts higher ratings for scale degrees 1, 4, and 5. The results show highest ratings for F and G in addition to those for C. However, ratings for E were also high, meaning that the three tones heard in the pattern, all members of the major triad, were judged as good completions and that this pattern of profile peaks is predictable by the tonal hierarchy model. These data in themselves are not decisive support for either model because the ratings for E do not differ significantly from those for F. The intervallic rivalry model might explain elevated ratings for the probe E, the only tone of these four that is not a possible tonic, in terms of its priming in the major triadic context itself; the high ratings for F, a probe that is a possible tonal center but that did not receive the benefit of priming in this pattern, are more impressive. With the tonal hierarchy model, the finding of higher ratings for the F probe might be interpreted as a chance statistical result.

In the diminished triadic context, interesting divergence from findings of Cuddy and Badertscher occurred with high ratings of C, the major-mode tonic. On average, this probe-tone rated 2.06, reflecting much stronger completion judgments than any of the other pitch classes within this context. This fits predictions of the intervallic rivalry model. The rating of 2.06 is only slightly lower than the average for pitch class 0 found in the other two contexts. However, it is important to observe that, in a real sense, the design of Experiment 1 is weighted against the intervallic rivalry model’s predictions for the diminished triadic context. Of the three, it is the only pattern that does not contain C. Both the major scalar and the major triadic patterns embed two instances of C in highly salient serial locations. Butler (1989, 1990b) has noted that any context tone might bias listeners in favor of it simply by virtue of repetition. In the case of the
tonic, both the major triad and the major scale contain two instances of C, whereas the diminished triad contains none. In light of this, it is all the more remarkable that listeners inferred the tonic from an otherwise neutral collection of pitches in the diminished triad. A second result consistent with predictions of the intervallic rivalry model is the modest peak at F#. In Experiment 1, diminished triadic tones occurred in an order less clear in pointing to C as good completion. In this ordering, members of the tritone switch roles, which increases the robustness of competing candidates, according to the intervallic rivalry model. In this case, the most obvious rival is F#, the keynote of F# minor (as illustrated in discussion of Figure 4b).

The data shown in Figure 7c share another similarity with those reported by Cuddy and Badertscher: Probe-tone profiles for the diminished triadic patterns are relatively flat when compared with profiles of the other two patterns. In spite of this, the correlation between profiles of Experiment 1 and those of Cuddy and Badertscher for this context was low ($r = -0.437, p < 0.156$).

Sensitivity to key, evidenced by high tonic ratings in the diminished triadic context, is accompanied by other influences on the profile shape. For example, diatonic tones are judged significantly better endings than are nondiatonic ones, and this suggests that ratings for diatonic tones not sounded can reliably reflect listeners' sense of "keyness." But several moderate "bumps" also appear in the probe-tone profile for this context and these, in principle, must be explained. It turns out that these bumps may be more consistent with the idea that listeners favor C major rather than F# minor. But the evidence for this turns, in part, on arguments promoted by the tonal hierarchy model because these smaller peaks occur at E and G.

For the intervallic rivalry model, neither of these tones has a claim to be a rival candidate for tonic, given this context. However, it remains possible that tonal reinforcing properties of the common intervals associated with E and G are at play in this context. Such factors are not explicitly developed in the rivalry model, which emphasizes the function of rare intervals. But if listeners are sensitive to salient pitches within the C major framework that reinforce the tonal hierarchy, then this would suggest that the diminished triadic context is providing specific information about key.

Another factor that might contribute to the bumps at E and G is melodic proximity. Note that in Figure 4a, the harmonic accompaniment allows the bracketed tones B and F to form the temporal succession of scale degree 7 to (scale degree) 4 (assuming the C major set). This succession resolves typically to the final melodic tone E, in the second measure, as the third of a presumed tonic triad. As such, it could provide a higher sense of melodic completion. Similarly, the slightly higher rating of G following the diminished triad (ordered as in Figure 4a) could be influenced by melodic and implied chordal factors within this key, such as the
step up from F to G, which “completes” an inferred V7 chord. The possibility that smooth melodic connections and proximity influence completion judgments complicates the interpretation of tonal profiles. Neither the intervallic rivalry model nor the tonal hierarchy approach readily accommodates melodic influences on completion judgments.

Finally, lack of an overall interaction of probe tone, pattern context, and training level (see Table 1) indicates that listeners of all three training levels responded similarly to the diminished triadic pattern. Thus, it is not likely that the reliance on either common or rare intervals is a tonal listening strategy limited to an elite group with high levels of musical training.

Experiment 2

Experiment 2, together with Experiment 1, provides a test of the temporal-order hypothesis of the intervallic rivalry model. Here, responses of listeners with different levels of musical training again yielded tonal profiles for three different melodic patterns. The contextual patterns in Experiment 2 differ in only one respect from those of Experiment 1: The same tones occur in each pattern, but in a new temporal order. Examples are shown in Figure 8. The major triadic arpeggiation (Figure 8a) now begins and ends with C, its root, while the major scalar pattern now can embed its two tonic pitches within randomized sequences (Figure 8b). The diminished triad (Figure 8c) represents the pitches of the tritone in an order optimal for C to be rated best completion according to the temporal-order hypothesis.

Experiment 1 yielded some data that provide clear support for the intervallic rivalry model in higher ratings for rival tonic candidates in the triadic contexts. But it also provided data consistent with the tonal hierarchy account, such as relatively high ratings for certain members of the C major triad in all contexts. In Experiment 2, temporal-order predictions of the intervallic rivalry model are pursued with the intent of resolving some of these differences. The primacy hypothesis addresses the temporal changes in the randomized major scalar patterns. Because all rearrangements of the scalar pattern continue to include the tritone, the intervallic rivalry model

![Fig. 8. Stimuli used in Experiment 2. One of the 12 randomizations of the C major scale is shown here. The probe tone was heard at the point corresponding to the rectangle in each example (a–c).](image-url)
predicts only modest changes in the tonal profile associated with this context in a completion task. More explicit predictions are possible for probe-tone judgments with the specific reordering of the two triadic patterns.

Consider first the major triadic context. The primary difference between the ordering of triadic tones in Experiment 2 and that of Experiment 1 is removal of the alternating root. One could argue that the regular alternation of the root (e.g., C-E-C-G) might have increased the likelihood of its rating as a good melodic completion. With a different arrangement (C-G-E-C), this alternation disappears in Experiment 2, although the tonic remains a repeated tone at highly prominent serial locations. Thus, the primacy hypothesis, applied to the major triadic patterns in both experiments, continues to predict an initial bias for C as a tonic. But the rare-interval hypothesis admits two additional keys as candidates for major tonal center for this pattern—F, and G. Furthermore, the rarest interval in these sequences—the major third from C to E—now appears last in this new temporal ordering, and this should intensify the possibility that the major triadic pattern could be heard as an arpeggiation of a dominant triad (i.e., instead of as a tonic triad). If so, the favored tonic should be F. Note that both features increase the likelihood that the pattern might sound like a better completion by moving to a tone other than the final one in the sequence. Extrapolations most likely from a tone heard last to completion in tonic judgments are the ascending semitone (here, D♭) and the descending perfect fifth (F) or its inversion, the ascending perfect fourth. In context with a C major triad, D♭, a smooth melodic continuation, is chromatic; F is diatonic in major keys of both C and F. In short, the intention of reordering the pitches of the major triad in Experiment 2 was to determine whether the probe tone F could be made a stronger rival for the tonic than in Experiment 1, where it was judged equivalent to E in resolution power.

A final goal in Experiment 2 was to evaluate the prediction of the temporal-order hypothesis that the reordered diminished triad should more clearly specify the major tonal center, pitch class 0, than in Experiment 1. In addition, because the rare-interval hypothesis recognizes that the major triad is inherently more ambiguous than the diminished triad, it also leads to the prediction that completion ratings for the tonic will be higher in the diminished triadic context than in the major triadic contexts, of either Experiment 1 and Experiment 2, for listeners at all levels of training.

METHOD

Subjects

Subjects were 45 adults who had not participated in Experiment 1, but who as a group represented a similar range of age and musical experience. The study group included 16
women and 29 men from 18 to 44 years old. The high group had a mean number of 9.5 years of performance lessons, the medium group, 8.3 years, and the low group, 6.2 years.

**Stimuli**

Temporal-order variations of the patterns used in Experiment 1 are notated in Figure 8. As in Experiment 1, both triadic patterns began on their roots, and comprised two roots, a third and a fifth; the variation was in the order of presentation of the second, third, and fourth tones in each triad. Order of tones in the major scalar patterns was a different randomized order for each probe tone in each block.

**Apparatus and Procedure**

The apparatus and procedure were the same as in Experiment 1.

**RESULTS**

The averaged probe-tone ratings for each pattern for each subject group of musical experience are shown in Table 2. There was a main effect of pattern \( F(2,84) = 62.93, p < .0001 \), with the mean rating for the major triad \( M = 3.63 \) higher than that for the diminished triad \( M = 3.74 \), which in turn was higher than that for the major scale \( M = 4.45 \). Probe tone also had a significant overall influence on ratings \( F(11,462) = 68.70, p < .0001 \). Musical experience groups performed similarly to each other and to the subject groups in Experiment 1. Analysis of normalized data

<table>
<thead>
<tr>
<th>Probe Tone</th>
<th>C</th>
<th>C#</th>
<th>D</th>
<th>D#</th>
<th>E</th>
<th>F</th>
<th>F#</th>
<th>G</th>
<th>G#</th>
<th>A</th>
<th>A#</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2.80</td>
<td>2.63</td>
<td>3.73</td>
<td>4.33</td>
<td>3.93</td>
<td>2.43</td>
<td>5.10</td>
<td>2.83</td>
<td>4.17</td>
<td>4.47</td>
<td>3.87</td>
<td>4.37</td>
</tr>
<tr>
<td>Medium</td>
<td>2.23</td>
<td>2.43</td>
<td>3.33</td>
<td>5.00</td>
<td>4.50</td>
<td>1.80</td>
<td>5.87</td>
<td>2.83</td>
<td>3.87</td>
<td>4.03</td>
<td>3.93</td>
<td>4.20</td>
</tr>
<tr>
<td>High</td>
<td>2.00</td>
<td>2.73</td>
<td>3.43</td>
<td>5.03</td>
<td>3.83</td>
<td>1.90</td>
<td>5.40</td>
<td>2.50</td>
<td>4.13</td>
<td>3.87</td>
<td>4.00</td>
<td>4.10</td>
</tr>
<tr>
<td><strong>Major Triad</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2.07</td>
<td>6.20</td>
<td>3.83</td>
<td>5.13</td>
<td>4.20</td>
<td>3.67</td>
<td>6.30</td>
<td>3.40</td>
<td>5.63</td>
<td>3.97</td>
<td>4.93</td>
<td>5.30</td>
</tr>
<tr>
<td>Medium</td>
<td>1.90</td>
<td>6.00</td>
<td>3.43</td>
<td>5.27</td>
<td>3.27</td>
<td>3.53</td>
<td>6.27</td>
<td>3.33</td>
<td>5.90</td>
<td>4.23</td>
<td>5.10</td>
<td>4.03</td>
</tr>
<tr>
<td>High</td>
<td>1.93</td>
<td>5.83</td>
<td>4.10</td>
<td>4.97</td>
<td>3.87</td>
<td>4.20</td>
<td>6.07</td>
<td>3.27</td>
<td>5.43</td>
<td>3.73</td>
<td>5.07</td>
<td>4.90</td>
</tr>
<tr>
<td><strong>Major Scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.57</td>
<td>3.60</td>
<td>3.97</td>
<td>4.40</td>
<td>3.67</td>
<td>4.00</td>
<td>3.97</td>
<td>4.27</td>
<td>4.97</td>
<td>3.13</td>
<td>3.67</td>
<td>3.10</td>
</tr>
<tr>
<td>Medium</td>
<td>1.37</td>
<td>4.17</td>
<td>3.60</td>
<td>4.73</td>
<td>2.90</td>
<td>4.47</td>
<td>4.67</td>
<td>4.43</td>
<td>4.83</td>
<td>3.40</td>
<td>4.77</td>
<td>3.53</td>
</tr>
<tr>
<td>High</td>
<td>1.63</td>
<td>3.57</td>
<td>3.50</td>
<td>4.90</td>
<td>3.10</td>
<td>4.17</td>
<td>4.07</td>
<td>3.83</td>
<td>4.60</td>
<td>2.40</td>
<td>4.17</td>
<td>3.47</td>
</tr>
</tbody>
</table>
showed that the interaction among pattern, probe tone, and musical experience was not significant \( F(44,924) = 1.04 \). Differences among profiles were supported by a significant interaction between pattern and probe tone \( F(22,924) = 20.90, p < .00001 \); it is this interaction to which attention is devoted.

Profile Analyses

As in Experiment 1, separate analyses of variance were performed on the profiles for each of the three pattern contexts. Profiles for the three contexts in Experiment 2 appear in Figure 9, along with profiles from Experiment 1 and from Cuddy and Badertscher’s experiment.

Major Scalar Pattern

The profile for the major scalar pattern in Experiment 2 appears in Figure 9b. Although it differs in some respects from that of Experiment 1, it remains highly correlated with the major scalar tone profile of that study \( r = .941, p < .0001 \). The correlation with the comparable profile of Cuddy and Badertscher is somewhat, but not significantly so, lower \( r = .804; Z = 1.325, p > .05 \).

The important features that scalar pattern profiles in Experiment 1 and 2 have in common are (1) the tonic C was rated higher than the dominant G \( F(1,42) = 36.54, p < .0001 \), thus indicating that removal of serial primacy for C did not challenge the relative sense of completion for the scalar pattern; (2) the tonic and dominant, taken together, were again rated higher than the mediant E \( F(1,42) = 35.60, p < .0001 \); (3) the pitches outlining the tonic triad were rated higher than other diatonic notes \( F(1,42) = 5.09, p < .05 \), suggesting a tonal hierarchy interpretation; and (4) finally, also consistent with a tonal hierarchy model, the presented scale tones were all rated higher than those not presented (chromatic tones) \( F(1,42) = 57.86, p < .0001 \).

The primary differences resulting from analyses of comparisons of judgments of individual probe tones were that two diatonic tones received higher ratings than anticipated on the basis of the Experiment 1 profile. (1) The probe corresponding to F (mean rating of 3.80) was rated as a significantly better completion than the three other nontonic-triad diatonic (D, A, B averaged) tones \( F(1,42) = 7.53, p < .01 \), but again ratings of F did not significantly differ from those of E (mean rating of 3.79); (2) neither E nor F was judged to be a significantly better completion than D (mean rating of 3.79); and (3) the probe corresponding to A was rated a somewhat better completion than B \( F(1,42) = 15.52, p < .0003 \).
Fig. 9. C major profiles from Experiment 1 (replication) and Experiment 2 (reordering), shown overlaid with profiles from Cuddy and Badertscher, 1987: (a) Profile in response to the C major triadic pattern, (b) profile in response to the C major scalar pattern, and (c) profile in response to the B diminished triadic pattern.
Major Triadic Pattern

Reordering tones in the arpeggiation produced a profile that was different from that recovered in Experiment 1. The most striking difference is that F outranked every other probe tone as the best completion. F received a mean rating of 2.04 whereas C received a rating of only 2.34. Listeners judged F to be a significantly better completion than E, which garnered a rating of 4.08 \( F(1,42) = 58.71, p < .0001 \). In fact, the context ratings of the median (E; mean rating of 4.08) indicated that it was heard as a slightly (but not significantly) worse completion than nontonic-triad diatonic tones (D, A, B; mean rating, 3.95). By contrast, the ratings of F, a tone not sounded in the contextual pattern, did not differ from those of C or G; all were equally good completions and superior to E. These statistics were reversed from Experiment 1 where, for example, F yielded a mean rating of 3.14, which did not differ from that for the E probe, while the ratings of C (1.67) indicated that it was judged the best completion in that context. Furthermore, in Experiment 2, combined ratings of all three C major triad pitches yielded a mean rating of 3.04, substantially below that of F. F was rated significantly higher than all other diatonic tones, save C and G \( F(1,42) = 90.63, p < .0001 \) and was the only diatonic tone not heard in the presented context to be so singled out. Such a remarkably high rating of F in this context cannot be anticipated in the tonal hierarchy model.

Other specific differences from Experiment 1 also appeared in the profiles. Completion ratings for C were lower than in Experiment 1 and did not differ significantly from those for G probes. Combined rating for triad members (C, E, G) did not differ significantly from those of other diatonic notes of the C major set (D, F, A, B). Here A, the keynote of the relative minor of C, was not judged to be a better completion than B, as it had been in Experiment 1. The chromatic tone, C\#'Db, received an unusually high completion rating (2.60) in this reordering, one that places it above a number of the diatonic tones, suggesting the influence of melodic proximity and bearing out the prediction of its influence on sense of completion.

Features of the two major triad profiles (Experiment 1 versus Experiment 2) that remain unchanged with temporal reordering are the following: (1) the probe pitches C and G, combined, were rated higher than E \( F(1,42) = 61.86, p < .0001 \); (2) the seven tones in the C major set were rated higher, as a set, than the tones outside it (C\#, D\#, F\#, G\#, A\#, or enharmonics) \( F(1,42) = 8.51, p < .01 \), although C\#, because of its ascending semitone relationship to the last tone of the stimulus, is distinctly higher; and (3) nondiatonic tones revealed similar pairwise differences, but again differences for the C\# probe were unusually marked (e.g., C\# was rated higher than D\# \( F(1,42) = 79.27, p < .0001 \) and C\# was
rated higher than the combination of D# with F# and G# \[F(1,42) = 175.15, p < .0001]\).

Figure 9a presents the major triad profiles of Experiment 2 along with those based on the other temporal orderings of the same pitches. Changes related to the reordering variable can be summarized in terms of correlations among these profiles. The major triadic pattern profile observed in Experiment 2 correlates poorly with that of Experiment 1 \((r = .541, p > .05)\). It also correlates relatively poorly with the profile reported by Cuddy and Badertscher \((r = .668, p < .025)\). The latter relationship, while significantly different from zero, nonetheless represents a significantly lower correlation with Cuddy and Badertscher’s data than in Experiment 1 \((Z = 2.350, p < .05)\). A large portion of the common variance between the two profiles in the latter case is attributable to the fact that, in both cases, the diatonic tones are rated as better completions than nondiatonic ones.

**Diminished Triadic Pattern**

As anticipated, the profile associated with this arpeggiation of triadic pitches is noticeably different from the profile of Experiment 1 (Figure 9c), as predicted. Overall, these two tone profiles remain positively correlated \((r = .626, p < .05)\), but the correlation between the Experiment 2 profile and that of Cuddy and Badertscher for the diminished triadic pattern is quite low \((r = .276)\) and nonsignificant \((p > .10)\).

The most important difference contributing to the changed profiles is that the tonic, rated similarly higher than all other probes, stands out more prominently. It was judged as a significantly better completion than the dominant \([F(1,42) = 129.11, p < .0001]\). In fact, the dominant, G, was not judged as an especially good completion tone in the reordered context. Nor was F judged a good completion; it was rated as a significantly poorer completion than other diatonic tones not contained in the tonic triad (D, A, B) \([F(1,42) = 14.50, p < .0005]\), a distinct difference with respect to its role in other contexts in both experiments. In short, the two viable candidates for the tonic, F and G, in the major triadic context, are not judged to be good completions in this context. Moreover, A was not rated significantly higher than B \([F(1,42) = 1.71]\). Another difference relates to F#: It no longer seemed to be a good completion, it was rated no better than the average of other chromatic probes (C#, D#, and G#) in this reordered context \([F(1,42) = 0.96]\). Other findings were: C# was rated higher than D# \([F(1,42) = 14.02, p < .0005]\), but C# was rated higher than D# together with F# and G# \([F(1,42) = 24.18, p < .0001]\).

Finally, there were two commonalities of the tone profiles for the diminished triad across Experiment 1 and 2: (1) tones diatonic to C major were rated higher than chromatic notes \([F(1,42) = 9.89, p < .01]\),
and (2) the mediant, E, received higher ratings than other diatonic tones (except C and A).

**DISCUSSION**

The data of Experiment 2 are consistent with predictions of the intervallic rivalry model and its temporal-order hypothesis. The most important result concerns the evidence that good completion can be controlled by orderings of tones in ways that influence key identification. This is evident in the tonal profiles for both triadic contexts. The specific diminished triadic arpeggiation produced the best completion for the tonic (1.52) in all contexts (and across both experiments). The mean score for the tonic probe in the diminished triadic context of Experiment 2 was markedly higher than that reported by Cuddy and Badertscher, suggesting that a factor contributing to their failure to find evidence for listeners’ use of rare intervals may have involved the ordering of pitches they used. In support of this is the low correlation of profiles in Experiment 2 with those reported by Cuddy and Badertscher. More recently, Cuddy (1991) reported an experiment based on a modified probe-tone task in which she found that even reordered versions of the diminished triad did not yield strong evidence for identification of the tonic. This finding is contrary to the results of Experiment 2. However, a potentially important difference between the two studies involves the fact that the reordered diminished triadic pattern in Cuddy’s experiment, relative to the major triad, also involved a shift of contour. The possibility of melodic influences on probe-tone profiles is considered below. Other negative evidence cited by Cuddy involved the lack of jagged tonal profiles for the diminished triadic pattern, the issue considered next.

In Experiment 2, the probe-tone profile for the diminished triadic context is not especially jagged. However, it does reveal two interesting features: (1) the major-mode tonic received a singularly high rating, and (2) diatonic pitches were judged better completions than nondiatonic ones. Having observed these features, it is important to point out that support for the intervallic rivalry model does not necessarily turn on the same jagged probe-tone profile as predicted by the tonal hierarchy model. Because the intervallic rivalry model is a model of key discovery, in principle, sufficient evidence for its support is singularly high ratings for a predicted tonic probe. The intervallic rivalry model implies that this pitch should be heard as the best resolution of a goal-oriented dominant-to-tonic harmonic motion set up by the F-to-B melodic tritone ordering. The fact that the tonic is missing in the diminished triadic context makes the presentation of this probe all the more salient to listeners as a completion and this contributes to the rationale underlying the temporal-order hypothesis. Ad-
ditional, more modest, peaks occur in the profile for the diminished triad at A. The explanation of the peak at A is straightforward in terms of the rivalry model: A is the relative minor of C major. The peak at E can be explained in terms of melodic factors, namely that E provides a sense of resolution by being the perfect fifth below the last tone heard as well as a good resolution of the melodic high point in the pattern (F to E). However, this explanation is technically outside the current scope of the intervallic rivalry model. At present, the rivalry model does not explicitly address melodic contributions to resolution, except insofar as they relate to orderings of the tritone interval (but see Brown, 1988).

Because the intervallic rivalry model implies that a single-peaked profile is evidence for lack of ambiguity in key identification, one of its strengths comes in addressing the absence of peaks associated with probes F and G. These are pitches that can function as potential tonics in the context of the major triad, but they can be eliminated as candidates for the tonic in the diminished triadic context by the flat profile over pitch classes 5, 6, and 7. This was not the case in Experiment 1, where the tritone ordering was not optimal for suggesting C major. In Experiment 2, the dominant of C major did not receive a high completion rating. If, in this temporal ordering, the diminished triad encourages listeners to infer harmonic motion from the dominant to the tonic, then a dominant pitch probe can be considered a continuation of this motion rather than a specific goal or resolution, in which event it should be rated an unsatisfactory completion. Similarly F, which received the highest rating in the reordered major triadic context, was heard as a relatively poor completion in the diminished triadic context. Thus, both of these pitches should be heard as part of a continuing harmonic motion, as part of a harmonic antecedent rather than a harmonic consequence. As such, they should be rated as poor completions. They were. Another peak in the data from Experiment 1 that was absent in Experiment 2 corresponds to a rival tonic, F# (pitch class 6). While F# was a viable completion in the ordering of the diminished triad in Experiment 1 (largely because that ordering set up dominant to tonic motion suggestive of the key of F# minor), in Experiment 2, this ambiguity and the corresponding peak vanished. This interpretation underscores the fact that the intervallic rivalry model cannot predict the same tonal profile as the tonal hierarchy model. In short, when taken as a whole, the tonal profile associated with the diminished triadic pattern, while not jagged, is nonetheless generally consistent with the idea that listeners are deriving a clear sense of key from the diminished triadic pattern when heard in this particular serial order.

With respect to the major triadic pattern, reordering of triadic pitches in Experiment 2 sought to evaluate whether the probe tone F could be considered a strong candidate for the tonic in this logically more ambiguous
context. If listeners do engage in position-finding strategies, then the intervallic rivalry model predicts that the three probe tones that will function as rival tonics are F, G, and C. Thus, all three should receive strong completion ratings in this context. This turned out to be the case. By contrast, the tonal hierarchy model predicts that C, G, and E should all be heard as good resolutions (completions) in this context, but it has no comparable rationale for F. However, the probe-tone judged to be the best completion, significantly better than E, in Experiment 2 was, in fact, F. Ratings for E, in this context, were much lower than for F and were comparable to those for several chromatic probe tones. These findings are consistent with predictions of the intervallic rivalry model: When F is heard after the C major triad, this succession can imply the harmonic progression of descending fifth as dominant-to-tonic in the key of F major or F minor just as logically as it can imply tonic-to-subdominant in the key of C major.

Finally, the randomized scalar patterns did not offer a basis for differentiating between predictions of the tonal hierarchy model and those of the intervallic rivalry model. But together these sequences do offer a referential context in which both models predict that clear tonal information is available. Although diatonic pitches offer substantially better completions than nondiatonic ones, only the probes corresponding to tonic and dominant pitches receive markedly better ratings among the diatonic tones. Again, the subdominant F figured as prominently as the mediant, E, and neither of these pitches stands out as offering distinctly better resolutions than the other diatonic tones presented in the context. The fact that listeners did not find E, a member of the tonic triad, to be a significantly more compelling completion than other nontonic-triad diatonic tones (e.g., F, D) is not in line with a strict interpretation of the tonal hierarchy model. In other respects, while reorderings may have contributed to slightly lower completion ratings of C, and slightly elevated judgments for F than in Experiment 1, these minor differences are outweighed by the general similarity between the scalar pattern profiles in Experiment 1 and 2. Overall, these data are consistent with the hypothesis that the major scalar pattern permits both a clear discovery of a tonal center via rare intervals and a reinforcement of the key via certain common intervals found within the scalar context. In sum, the resulting tonal profile provides support for both intervallic rivalry and tonal hierarchy models.

General Discussion

Results of Experiment 1 and Experiment 2 together support predictions of the intervallic rivalry model about key discovery. The most relevant predictions are associated with the temporal-order hypothesis, which main-
tains that whenever a musical event contains a tritone arranged so as to imply a dominant-to-tonic harmonic succession, this configuration will lead to clear key discovery.

Three specific bases for these and related predictions of the intervallic rivalry model are found in the two experiments. First, with respect to the rare-interval hypothesis, the intervallic rivalry model predicts that the diminished triad provides unambiguous major mode information (barring chromaticism) about key. This hypothesis finds support in the fact that listeners in both Experiments 1 and 2 rated the tonic probe to be the best completion for the diminished triadic context, in spite of the fact that this pitch did not appear in the presented patterns. Second, the rare-interval hypothesis also predicts that the major triadic context offers information that is more ambiguous tonally than the other two contexts. Consistent with this analysis is the finding that listeners in both Experiments 1 and 2, as well as the original experiment by Cuddy and Badertscher, rated the probe pitches C, F, and G (tonics of the three major keys in which a C major triad is diatonic) as the best completions for the major triadic patterns; all three were judged as significantly better completions than other probes in Experiment 2. Third, the temporal-order hypothesis addresses the temporal ordering of events within the diminished triadic contexts across the two experiments. In Experiment 2, where the rarest-interval rearrangement was predicted to optimize key identification in C major, the diminished triad produced a relatively flat tonal profile in which ratings of the tonic probe alone stood out as reflecting the strongest sense of completion reported across all six conditions studied (Experiments 1 and 2). It seems clear that the rare-interval hypothesis, together with the temporal-order hypothesis, finds support in these data.

Perhaps most critical for the intervallic rivalry model are comparisons between the two kinds of triadic contexts (major versus diminished). Whereas the major triad, in the abstract, is ambiguous with respect to major-mode key identification, the diminished triad is not. In terms of the findings, the most striking difference between these two contexts is that the major triad produces distinctly jagged peaks in its probe tone profile whereas the diminished triad tends to approximate a single peaked function. Here lies the connection between theory and data: It is precisely because the major triad can be heard in multiple keys that a jagged tonal profile occurs. According to the rivalry model, possible candidates for the tonic such as (here) C, F, and G may, at one time or another, influence completion judgments in this context. Consistent with this view, it turns out that all three of these pitches were rated as strong completions, and in Experiment 2, the profile peak for the F probe even exceeds that for C and G, the two probes heard in the patterns. This suggests that a significant percentage of listeners may have heard the
major triad in the key of F, not C. The issue of good completion ratings for the mediant in the major triad context is relevant here, because the intervallic rivalry model does not predict those whereas the tonal hierarchy model does. In Experiment 1, E was rated as no better a completion than F, whereas in Experiment 2, E was judged to be a relatively poor completion (relative to all three rival tonics). If stability within a tonal hierarchy were the primary determinant of judgments in Experiment 2, then ratings of E should be significantly higher than those of F in both experiments. They were not. All of these findings suggest that multiple peaks arise, in part, as a result of key ambiguity of the related intervals and that the peaks reflect listeners’ response to rival tonic candidates.

Conversely, the intervallic rivalry model does not predict a jagged profile for triadic contexts that are not ambiguous. To the degree that such a context specifies a single pitch as tonic, the probe-tone profile should reflect the listener’s certainty of the single best completion with a single-peaked probe-tone profile. This leads to the prediction that the tonal profile for the diminished triad, optimally ordered, should be relatively flat, as it was.

This discussion implies that probe-tone profiles can be interpreted in terms of multiple implications of pitch intervals within a presented context. But if this is the case, how can we explain the jagged tonal profile that typically appears when scale patterns provide the context for probe judgments. Do these profiles also reflect an ambiguity in this context? We have maintained that the scale pattern, finally, is not ambiguous with respect to key identification, because it does contain rare intervals. Why then does a jagged probe-tone profile appear with this pattern? In part, the answer may rest with the notion of implication. The complete scale pattern also contains pitches and implied chords that figure in related keys, and so it does imply related keys. However, in this context, the strongest implication is the key associated with the discovered tonic. That is, in a context in which both rare and common intervals are displayed, the key is not only clearly specified but is reinforced by implications of the more common intervals that connect it to related keys. Here probe pitches associated with common intervals presented in a tonal context, where they are clearly linked to a specified keynote, should be judged to fit better than others because in this case their multiple participation in other keys reinforces the discovered key (e.g., pitches a third, a perfect fourth or perfect fifth away from pitch class 0).

Although this interpretation does not emphasize the intrinsic consonance of specific diatonic pitches that reinforce tonality within an established key, it is similar in other respects to interpretations made by Krumhansl and her colleagues (e.g., Krumhansl, 1990a; Krumhansl & Kessler, 1982). That is, the interpretation we place on multiple peaks in
scalar probe-tone profile is that they can reflect, among other things, a
reinforcing function of common intervals within the diatonic set, once the
correct keynote of the set has been discovered. Once a rare interval has
served its function and permitted the listener to determine an unambigu-
ous key, tonal functions related to common intervals and their implica-
tions with respect to related keys can influence completion judgments. It is
these latter relationships, which involve common intervals, upon which
the tonal hierarchy capitalizes and that it describes so well. Effectively, the
intervallic rivalry model and the tonal hierarchy model complement one
another insofar as the former may function with rare intervals to guide
and confirm key discovery, while the latter may come into play with
common intervals to express tonal implications between the established
key and related ones.

It is clear that the potential of the probe-tone profile as a testing ground
for different theoretical interpretation has been enlarged. We have sug-
gested that the shape of a profile may be sensitive not only to ambiguity or
multiple implications of intervals in different keys but also to harmonic
motion (continuation versus completion). Clearly, current theories empha-
size different determinants of these profiles. But along with the potential for
exploring different determinants of probe-tone profiles, as such, come some
complications. Good completion or its converse, good continuation, can
arise from multiple determinants and if these various influences on listeners’
completion judgments cannot be identified in advance, they may baffle and/
or confound a clear interpretation of the peaks in a tonal profile.

In the present research, one illustration of this involves the effects of
melodic completion (or continuation) completion ratings. We cite three
examples. In Experiment 1, the alternation of C in the major triad, C-E-C-G,
could lead the listener to expect another C to follow and thus the probe
that confirms this melodic expectancy might receive a higher completion
rating. Second, in Experiment 1, peaks associated with probes E and G in
the diminished triadic context could be interpreted both as favoring predic-
tions of a tonal hierarchy model or in terms of melodic proximity conferred
by voice-leading from the pattern to the particular probes in that context.
The fact that the peak associated with G vanished in Experiment 2 for this
context suggests that melodic proximity may have been a contributing
factor. On the other hand, in Experiment 2, the high ratings of Db (pc 1) in
the major triadic context (see Figure 9a) are easily interpreted in terms of
melodic completion. In Experiment 1, the Db presents an awkward melodic
leap of a tritone from the last tone heard in the contextual pattern and so
could conceivably receive a weak completion rating for this reason rather
than because it is low on the tonal hierarchy. In Experiment 2, Db provides a
smooth connection of an ascending semitone, which contributed to its
much higher relative and actual rating, a result that cannot be attributed
simply to pitch proximity. Semitone motion is particularly suggestive of both continuation and completion in tonal melodies as part of the resolution from dominant to tonic harmonic function. The ascending semitone motion from the end of the contextual pattern to the probe tone Db is interpretable as a small-scale reference to leading tone to tonic. Descending semitone motion in major mode is often a reference to subdominant to mediant. (These common tonal patterns are notated in Figure 4.)

A second cautionary note with regard to interpretations of probe-tone profiles comes with possible confoundings associated with context manipulations. By necessity, the presentation of a tonal context of any sort in the probe-tone procedure has the potential for biasing listeners in ways entirely unrelated to the models under consideration here. For example, a simple memory explanation would maintain that listeners should respond to pitches presented in a previously unsounded context in a manner systematically different than to ones not-so-recently presented (see also Butler 1989, 1990b). Thus, one interpretation of the jagged tonal profile found with the major triadic context simply attributes it to the influence of tonal repetitions on memory. The finding that E is judged a better completion than other (not sounded) nontonic-triad diatonic tones in the major triadic context—but not in the major scale context (in both experiments)—may be related to this confounding. Although this factor cannot be entirely ruled out, an analysis by Cuddy (1991) suggests that it is not the whole story. Repeated presentations of certain context tones may indeed bias listeners toward them as probes, and this probably does stack the deck against the intervallic rivalry model. But there are other indications in these data that listeners are not simply relying on memory for repeated tones. The clearest evidence, of course, comes from the very strong completion judgments of the unsounded tonic in the diminished triad context, where at the same time no significant peaks were associated with tones actually present in that context. In addition, the fact that F, a tone not sounded in the C major triad, received very strong completion ratings is also indicative of the fact that people are attempting to make tonal inferences in ways not strictly determined by tone repetitions. Thus, although we cannot rule out the possibility that tone repetition effects account for some of the completion rating score in some conditions, it seems unlikely that repetition effects account for the entire pattern of data observed.

These sorts of considerations suggest that from a listener’s perspective, instructions to judge the “goodness of completion” of a probe tone in some context may, itself, be ambiguous, which, in turn, leads to several observations. First, it reinforces the point that probe-tone profiles do offer significant potential for pinpointing various determinants of completion, provided confoundings can be avoided and clear predictions formulated. Second, it suggests the need for measures addressed to disambiguating the
task: More specific instructions or manipulation of instructions may be helpful. As the cautionary comments by Butler (1989, 1990b) suggest, completion judgments may reflect the influence of several converging factors. This underscores the argument that experimental evidence for key discovery must finally incorporate data from other tasks as well. In the present experiments, a completion task was used to address criticisms of the intervallic rivalry model arising from experiments of Cuddy and Badertscher. But the present results suggest that in the future completion tasks should be supplemented with more specific instructions regarding key identification or with data from a converging task. When, for example, Brown and Butler (1981) used an explicit key-identification task, they reported results that converge with those of Experiment 2—listeners were most accurate picking the true tonic when they were provided with the most tonally specific context.

Finally, the fact that performance across these various contexts produced much the same pattern of results for all levels of listeners suggests that key discovery and responsiveness to tonal implications of rare and common intervals are not the preserve of specially trained musicians. Indeed, it is essential that the average listener who routinely enjoys tonal music be able quickly and effortlessly to discover a cognitive pitch anchor for the unfolding musical surface. The intervallic rivalry model suggests one way that this process may occur.3

References

Butler, D. Describing the perception of tonality in music: A critique of the tonal hierarchy

3. Portions of this study have been presented (Brown & Butler, 1988) and reported (Brown & Butler, 1989). The authors are grateful to Michael Lynch for thoughtful comments on the manuscript.