Psychol. Belg.
1991, XXXI-2, 197-216.

REPRESENTATIVENESS OF DIATONIC STRUCTURE AND DETECTION OF OUT-OF-SCALE NOTES

Robert J. WEST
Department of Psychology
Royal Holloway and Bedford New College
University of London

Ian R.M. CROSS
Faculty of Music
University of Cambridge

Peter HOWELL
Department of Psychology
University College London

This paper describes four experiments investigating the detection of out-of-scale notes following brief, diatonic-conformant context sequences. Subjects were presented with diatonic scale-conformant three-note sequences of different kinds and were asked after each one to indicate to what extent subsequent probe notes were heard as "fitting". Listeners more clearly identified out-of-scale notes following sequences comprised of pitch-class sets which can be formed in many places within the diatonic scale — such as F-C-D or F-G-D — than they could when the context sequences consisted of pitch-class sets which can occur in fewer places in the diatonic scale, such as F-G-C. The results supported the hypothesis that the extent to which context sequences were representative of diatonic scale structure influenced listeners' judgments.

INTRODUCTION

A musical scale is a limited set of pitches used in musical composition. Almost all Western music since 1600 restricts the use of pitch to twelve different pitches per octave. In the current equal temperament system of tuning these twelve notes are conceived of as dividing the octave into twelve exactly equal frequency-ratio steps, constituting the chromatic scale. Most scales identified in music theory make use of only seven of these notes, and the most fundamental scalar arrangement of seven different notes out of the twelve notes of the chromatic scale is called the *diatonic scale* (corresponding to the white notes on a piano).

The diatonic scale can be thought of as a particular *pattern of intervals*. While we often talk of "the scale of C major", what we are usually referring to is one of two things: either a stereotypical ascending and descending linear pattern of notes, starting on a note C and finishing on a different (octave-related) note C and making use of only the white notes on a piano keyboard, or a type of organisation underlying a particular piece of music, characterisable by the fact that the note C functions as a "tonal centre" — a stable reference point — within the piece, and that only the white notes are used. In this latter sense, we should more accurately say that the piece is in the *key* — rather than the scale — of C. In fact, the idea of scale — as opposed to key or mode (see Dowling, 1978) — bears no implication that one note of the scale, or *scale degree*, is in any sense more important that any other, but implies only the existence of a collection of notes, which is a restricted subset of all of the notes available within the chromatic scale, and between which certain patterns of intervals may be formed. The idea of the diatonic scale as a structural entity of this type is probably most explicitly expressible in terms of a group-theoretic account of musical pitch.

In his group-theoretic account of musical pitch, Forte (1973) treats the notes of the chromatic scale as analogous to the set of numbers 1 to 12, or 0 to 11. Within this approach, all octave-related notes are treated as functionally identical. This enables the notes of the chromatic scale — or chromatic *set* — to be depicted in two different circular, or cyclical, forms; these are directly analogous to the chroma circle and to the circle of fifths. This approach also relies on the idea that sets of notes, or sets of *pitch classes*, can be regarded as functionally identical if one pitch-class set can be transformed into another by means of a simple operation such as transposition or a compound operation such as transposition together within inversion. Sets of pitch classes are also regarded as functionally identical if they have the same membership, irrespective of the order in which the members of the pitch-class set are laid out. To give an example, the set (2,0,7) — or D-C-G — would be regarded as identical to the set (0,2,7) — or C-D-G — and to the set (2,4,9) (D-E-A). In other words, within a group-theoretic representation of musical pitch, two sets of notes or pitch classes are identical when the same structural relations can be shown to hold for each set. This is somewhat of an over-simplification, but will suffice as a definition for the present purpose.

Balzano (1980, 1982) points out that within a group-theoretic representation, the pitch-class set — or set of notes — corresponding to the notes of the diatonic scale has certain apparently unique properties; it has the property of *uniqueness*, in that each pitch class can be differentiated from every other pitch class by the set of intervals that it forms with each other pitch class of the set: it has the property of *simplicity*, in that each diatonic scale or diatonic set may be transformed into another such set by changing only one pitch class, the relation between the initial set and the transforming element being the same for all equivalent forms of the diatonic set: and it has the property of *coherence* in that the sum of any two intervals formed between three adjacent pitch classes or notes of the set will be greater than any single interval occurring between adjacent pitch classes. Coherence and uniqueness can be thought of as conferring possible advantages in perception, in that the notes of a diatonic melody will be differentiable simply by the intervals that they form with other notes irrespective of their order of occurrence and any major or minor interval within the melody will be constituted of the same number of scale steps enabling a listener to judge the size of movement that it constitutes within the underlying scale (with the sole exception of the tritone). The diatonic set is the only set within a group-theoretic representation of musical pitch which exhibits these particular properties.

Possible implications of these properties of the diatonic set for perception have been further explored by Browne (1981). He suggests that diatonic set structure uniquely facilitates cognitive activities which he describes as *position finding* and *pattern matching*. Position finding involves questions concerning orientation and reference; questions such as "where are we" (in respect of tonal space) and of "how long should I track what happened in terms of note x" are characteristic of position finding processes. So the identification of the tonic (or other structural notes) within a passage or the tracking of modulations between tonal regions could both be described as position finding processes. Pattern matching involves questions of similarity and belongingness, such as "is this the same as that", or "does this belong to the same thing as that".

Browne points out that within the diatonic set, intervals and subsets of the diatonic set — a diatonic subset is a set formable between any members of the diatonic set — can be regarded as either more rare or more common. Thus the interval of a semitone occurs in only two positions within the set, while the interval of perfect fourth occurs in six positions. Similarly the pitch-class set (0,2,6) — analogous to the set of notes F, G and B — can only occur in the one position within the set, while the pitch-class set (0,2,5) — analogous to the set of notes C, D and F — can occur in eight different positions (assuming that equivalence between pitch-class sets is on the basis of transposition together with inversion). Browne suggests that rare intervals may subserve functions in cognition involving differentiation between notes within the diatonic set so as to derive a tonal hierarchy (position finding), while common intervals may subserve cognitive functions involving association and integration of notes as members of the set (pattern matching). Thus Browne suggests that the structure of the diatonic set may play a dynamic role in the cognition of musical pitch and pitch relations.

The importance of diatonic structure in the perception and evaluation of Western music is well recognised. Diatonic structure markedly influences memory for, and appreciation of, music (Deutsch, 1982; Cross, Howell & West, 1983). What is less clear is how a sense of diatonic conformance (i.e., an impression that a given set of notes is derived from the diatonic set) is evoked in a listener. This paper seeks to clarify this by investigating how information provided by the pitches of three in-scale notes played in sequence can enable a subsequent out-of-scale note to be perceived as poorly fitting.

Krumhansl and others have shown that when listeners hear the notes of the diatonic scale in ascending or descending order, or when they hear a major triad (e.g., the notes of the chord C-E-G), they perceive subsequent out-of-scale "probe"notes as musically unstable and the indicated tonics as "stable" (see e.g., Krumhansl & Shepard, 1979; Krumhansl, 1983; Krumhansl, 1990; Cuddy & Badertscher, 1987).

We have examined the issue in a slightly different way — looking at the use of subsets of the diatonic scale, and using a variety of different temporal orderings of notes in sequence to balance factors that might otherwise contribute to the cuing of a "tonal hierarchy" — or sense of one note being more "stable" than others — by sensitivity to the structure of an "event hierarchy", such as the "end-primacy" suggested by Deutsch (1984) and Bharucha (1984). The purpose was to identify to what extent *set membership* of context notes would play a role in helping to identify subsequent out-of-scale probes. In a previous series of experiments, we found that subjects' abilities to detect out-of-scale notes were a positive function of the "logical diatonic scalar ambiguity" of the preceding notes (Howell, West & Cross, 1984); i.e., our subjects formed a stronger impression of diatonicism from within-scale note sets that were *more* rather than *less* ambiguous as to the diatonic scale from which they were drawn, irrespective of the order in which the context notes were presented. For example, F# was easier to detect as ill-fitting in sequences such as F-C-A-G-F#-D (or indeed G-A-F-C-F#-D) than in sequences such as F-C-B-G-F#-D. In the first sequences, the notes preceding the F# could have been drawn from three different diatonic major scales (on C, F and B*b*), while the notes preceding the F# in the second sequence were scale-specific in that they could only have come from the diatonic major scale on C. In a subsequent study (Cross, Howell & West, 1985), we also found that listeners judged out-of-scale probes to fit less well with maximally ambiguous three-note sets (e.g., F-C-G) than with maximally scale-specific sets (e.g., F-C-B).

These findings appear to conflict with studies by Brown and Butler (e.g., Brown, 1988; Brown & Butler, 1981) which showed that three-note sequences containing the tritone (thus deriving from only one possible diatonic scale, and being completely unambiguous) were highly informative as regards tonal centre. For example, listeners were more likely to produce C as the tonal centre following a context sequence of F-G-B than, say, F-G-C. However, ours and Brown's tasks were different in a way which may prove to be important. We were concerned with listeners' judgments about membership of the diatonic set, whereas Brown was concerned with identification of tonal centre. Cuddy and Badertscher (1987) found that the diminished triad — the set (0, 3, 6,) containing the tritone and therefore completely *unambiguous* as to the diatonic scale from which it is drawn — was very unhelpful in allowing subjects to discriminate between diatonic and non-diatonic probe tones.

In our previous experiments it seemed implausible that diatonic scalar ambiguity per se would be the main variable of interest. A number of other hypotheses may explain the result.

The first possible explanation is that the crucial variable is the commonness of occurrence of the different sequence — more properly, *set* — types in the listeners' musical experience. Thus, highly scale-ambiguous sets of notes such as F-C-G (which can occur in five different diatonic scales, irrespective of order of note occurrence) probably occur more in Western tonal music than do sets such as F-C-B (see Simonton, 1984). If the former are more characteristic of tonal music, they may make it easier to identify subsequent deviation from diatonic structure (although tonality and scale structure are by no means the same thing, see Cross, West & Howell 1991).

Secondly, it may be that the smaller the number of possible out-of-scale notes, the greater might be their salience in that context. Thus sequences such as F-C-G only have one note out of scale with them (F#), while sequences such as F-C-B have five possible out-of-scale notes (F#, C#, G#, E*b*, B*b*).

Thirdly, out-of-scale notes will in general be further away in terms of fifths from scale-ambiguous sequences (Figure 1). For example, the out-of-scale note f# is further away in terms of fifths from the members of the scale-ambiguous set, F-C-G, than from the members of the scale-specific set F-C-B. If proximity in the circle of fifths is one dimension of judgments of similarity of pitch-classes (e.g., Krumhansl & Shepard, 1979), then one could clearly expect the F# to be rated as more ill-fitting with the F-C-G type of sequence that the F-C-B type of sequence.

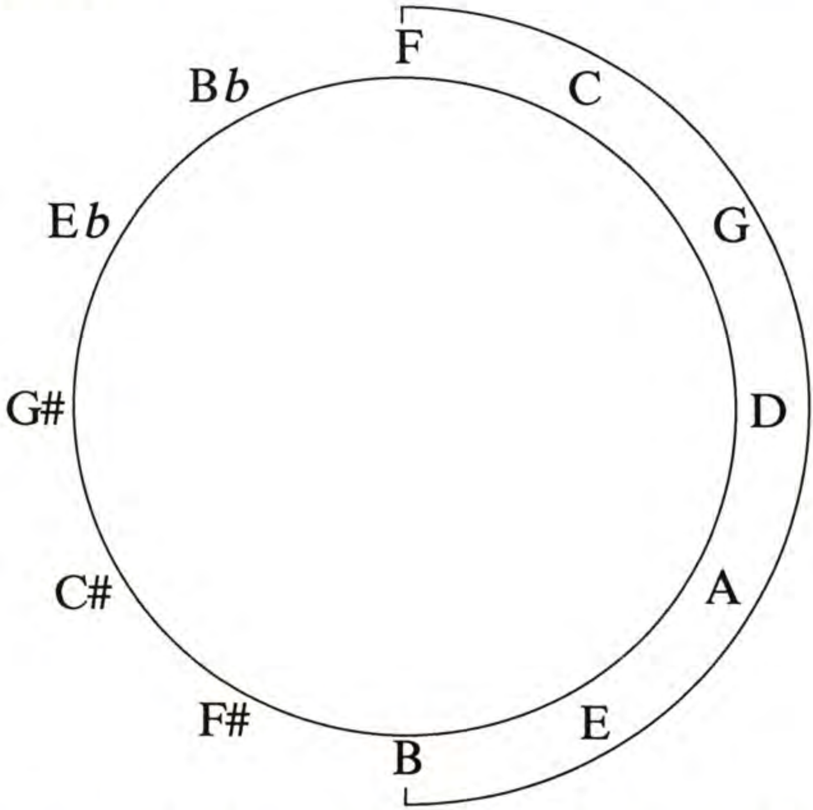

Figure 1: Diatonic scales are connected regions of seven notes within the circle of fifths (e.g., F clockwise through B gives the scale of C). The circle of fifths representation makes it easy to determine how "scale-ambiguous" a particular note set is. Because the notes F and B (a tritone apart) are at the extremes of the seven-note region in the circle of fifths any set of three or more notes from the scale on C containing these notes is logically unambiguous as to the scale from which it could have been drawn.

Fourthly, the critical factor may be how far the intervals in a given sequence type are representative of diatonic scale structure in the abstract. Thus the set of intervals formable within F-C-G occurs at five different places in the diatonic scale on C (i.e., F-C-G, C-G-D, G-D-A, D-A-E, A-E-B), giving it a diatonic multiplicity of five, whereas the set of intervals in F-C-B can occur in only one position and thus has a multiplicity of one. This is easily seen in Figure 1 where the two notes F and B fall at the margins of the seven-note set which is the diatonic scale on C; there is no way in which this pattern of notes (intervals) can be strictly transposed so as to remain within the scale on C. Thus, disregarding any biases resulting from familiarity with aspects of prevailing musical style such as the statistical distributions of occurrence of sets (as suggested in the first hypothesis), the set-type F-C-G is more likely to occur within music that is diatonic than is the set-type F-C-B. This hypothesis essentially extends Browne's (1981) suggestion that common diatonic intervals are likely to subserve a function in cognition in pattern matching activities to embrace the notion that the diatonic commonality of pitch-class sets may be similarly functional in perception.

This raises the question of what three-note set should be construed as *most* "representative" of diatonic structure. In assessing the formal diatonic representativeness of a given pitch-class set we have to determine the criteria on which to "reduce" sets of pitch classes to a single set-type, as this will determine the multiplicity of the set-type within the diatonic set. As indicated previously, sets of pitch classes can be regarded as functionally identical if one pitch-class set can be transformed into another *either* by means of a single operation such as transposition *or* by means of a compound operation such as transposition together within inversion (see Balzano, 1982). Moreover, it should be noted that in conceiving of these three-note sequences as pitch-class sets we are treating *ordered* sequences of pitches as though they were exemplars of *unordered* pitch-class sets. In doing this we are implicitly hypothesising that the interval pair $+7$-5 between the notes F-C-G will share perceptual qualities with the interval pair $-7+5$ between the notes G-C-F.

If we just consider how many different transpositions of a three-note set are possible while remaining within one diatonic scale or set, then F-C-G-type sets are most representative. These can be transposed five times and still remain within diatonic set (they have a transpositional multiplicity of five): F-C-D and D-F-G-type sets have a transpositional multiplicity of four: F-C-A, F-G-A and F-D-A-type sets have a transpositional multiplicity of three, and so on. However, it is important to remember that it is the pitch interval patterns or pitch-class set types, that are relevant here, not the particular pitch classes used. This means — as is implicit in the above idea of sets as *unordered* — that we can take into account the fact that intervals appear to share perceptual qualities whether they go up or down. Thus an interval of -7 semitones may be equivalent to one of +7 semitones. It may also be the case that intervals which result in a movement to the same pitch class irrespective of octave — which are *inversionally equivalent* — are equivalent in perception. Thus moving from C to the G above involves an interval of +7 semitones; moving from C to the G below involves an interval of -5 semitones. So +7 and -5 are complementarily, or inversionally, equivalent.

If the idea of inversional (complementary) interval equivalence holds true in perceptual judgments about diatonic structure — that is, if equivalence between sets is based not solely on transposition but on transposition together within inversion — then it can be shown that the set-type which is most representative of diatonic structure is *not* the F-C-G-type but rather the F-C-D-type, which has a transpositional-inversional multiplicity of eight. This can be seen most clearly by noting that when inverted and transposed the set F-C-D becomes the set D-G-F; as both of these types of sets have a transpositional diatonic multiplicity of four, when "reduced" to the same set-type that set has a transpositional-inversional diatonic multiplicity of twice four, or eight (see Table 1).

Table 1. — *Representativeness Scores of Sequence Types Used in Experiments*

Sequence type	Reflection	Score based on transpositions only	Score based on transpositions and reflections
F-C-G	F-C-G	5	5
F-C-D	F-G-D	4	8
F-G-D	F-C-D	4	8
F-C-A	F-D-A	3	6
F-D-A	F-C-A	3	6

In fact, the degree of representativeness of a three-note sequence as defined above can be worked out by first assessing whether the representation of the set within the circle of fifths has a reflection about its midpoint which is different from the original. F-C-G reflected around its midpoint (C) is still F-C-G: F-C-D reflected around its midpoint (the space between C and G) gives F-G-D: F-C-A reflected around its midpoint (G) is F-D-A, and so on (Figure 2). If this reflection process results in a three-note set which is different in content from the original we should consider the two reflections as equivalent and double their (transpositional) representativeness score. When the set is inversionally symmetrical — such as F-C-G, which does not yield a different set under reflection — its transpositional-inversional representativeness score remains the same as its transpositional representativeness score.

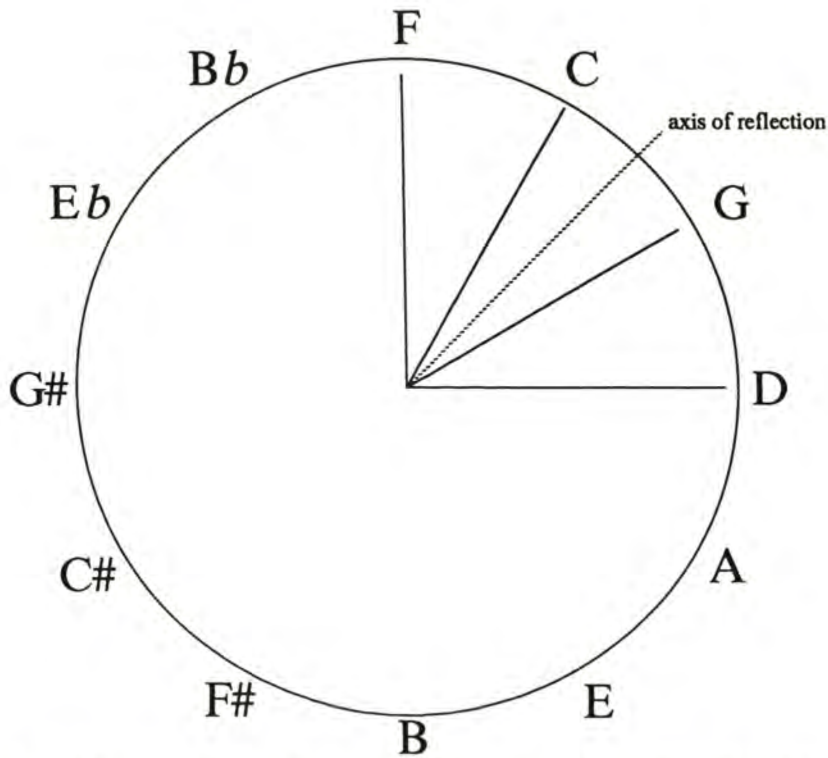

Figure 2: Note sets which are reflections of other note sets across the midline of their circle of fifths representation have the same set of intervals between elements in all their various orderings, i.e., are exemplars of the same type of pitch-class set. The three-note sets F-C-D- and F-G-D are reflections of each other in this way, and hence fall into the same pitch-class set type (here, the pitch-class set 0,2,5). The note sets F-C-A and F-D-A are also equivalent in this respect (constituting the pitch-class set 0,3,7) as are G-D-E and G-A-E (again, pitch-class set 0,2,5), etc.

We set out to determine which of these four hypotheses was correct. The first three hypotheses all lead to the prediction that listeners will obtain a better sense of diatonic structure from F-C-G-type sequences than from any other three-note set. If the fourth hypothesis is reformulated so as to take into consideration the possibility that diatonic representativeness is based on transpositional-inversional multiplicity, than it predicts that F-C-D/F-G-D-type sequences will lead to a better sense of diatonic conformance. As we were interested in the relationship between pitch-class set-type and sense of diatonic conformance our experiments used many different randomly selected orderings of the three notes in the context sequences. This is not intended to suggest that temporal ordering is of no importance in the perception of tonal relations (as Brown, 1988, gives evidence to the contrary), but that it is not significant in determining the nature of the sensitivities displayed by listeners to atemporal, diatonic structure in the abstract.

Experiments 1 and 2

The first two experiments used a variant of what Krumhansl (see Krumhansl, 1990) has called the "probe-tone" technique. Subjects listened to three-note sequences and for each one indicated how well they thought a fourth, probe, note fitted in with the previous three notes.

We examined ratings of fittingness of out-of-scale probes following F-C-D/F-G-D-type sets compared with ratings of out-of-scale probes following F-C-G-type sets. The experiments were near exact replications but on different subject samples, one with little or no musical training and the other made up of undergraduate and postgraduate music students. If the F-C-D/F-G-D-type of sequence made the out-of-scale probe seem to fit less well than with the F-C-G-type, this would militate against any but the last of the hypotheses outlined earlier. It would support the idea that it was how far the intervals embodied in the three-note sets (irrespective of interval order) were representative of diatonic scale structure which was important.

Subjects and methods

The subjects for Experiment 1 were five male and seven female undergraduate and postgraduate students at University College London. Most had received some tuition on musical instruments at an early age, but none had taken their study of music beyond elementary level. The subjects for Experiment 2 were nine undergraduate and postgraduate music students at The City University, London. Five of these were male.

The sequences were generated and played on a PDP-11 mini-computer, the sound having been synthesised on a Fairlight CMI as a fundamental and its first two odd harmonics and recorded to the PDP-11 at a 10 KHz sampling rate through a low-pass filter at 4.5 KHz (48 dB/8ve). The notes used ranged upwards for two-and-a-half octaves from E below middle C. The subjects sat in sound-attenuated rooms and listened to the sequences through headphones. They were instructed to listen to each sequence and rate how well they thought the fourth note fitted in with the preceding three. The notes were played at a rate

Table 2. — *Examples of Three-note Sets Used in the Experiments*

Note set	Possible keys	Possible out-of-scale notes
F-C-G type sequences (transpositions of F-C-G)		
F-C-G	C,F,B*b*,E*b*,G#	F#
C-G-D	G,C,F,B*b*,E*b*	C#
G-D-A	D,G,C,F,B*b*	G#
D-A-E	A,D,G,C,F	E*b*
A-E-B	E,A,D,G,C	B*b*
E-B-F#	B,E,A,D,G	F
etc.		
F-C-D/F-G-D type sequences (transpositions of F-C-D/F-G-D)		
F-C-D	C,F,B*b*,E*b*	F#,C#
F-G-D	C,F,B*b*,E*b*	F#,C#
C-G-A	G,C,F,B*b*	C#,G#
C-D-A	G,C,F,B*b*	C#,G#
G-D-E	D,G,C,F	G#,E*b*
G-A-E	D,G,C,F	G#,E*b*
etc.		
F-C-A/F-D-A type sequences (transpositions of F-C-A-/F-D-A)		
F-C-A	C,F,B*b*	F#,C#,G#
F-D-A	C,F,B*b*	F#,C#,G#
C-G-E	G,C,F	C#,G#,E*b*
C-A-E	G,C,F	C#,G#,E*b*
G-D-B	D,G,C	G#,E*b*,B*b*
G-E-B	D,G,C	G#,Eb,B*b*
etc.		

Note. The actual order of presentation of notes in any given three-note sequence was randomly determined.

of two per second with a six second pause between each four-note sequence for subjects to make their ratings.

There were two conditions, corresponding to the use of two three-note sequence types. These were random transpositions of (i) F-C-G and (ii) F-C-D/F-G-D. The probe was always an equivalent transposition of F# (the only note which does not fit into the same diatonic set as F-C-G as well as F-C-D/F-G-D). The ordering of the notes in each set was randomised. Each subject listened to 12 sequences of each type. In Experiment 1 the subjects rated fittingness on a seven-point scale, while in Experiment 2 a five-point scale was used. Examples of sequences and probes used in both experiments are given in Table 2.

Results

The results of both experiments indicated that out-of-scale probes were judged less well fitting after the the F-C-D/F-G-D-type than the F-C-G-type. In Experiment 1 the means were 3.8 versus 4.4 respectively, $t(11)=3.6$, $p<0.005$. In Experiment 2 the corresponding means were 2.3 and 2.8, $t(8)=3.2$, $p<0.05$. It appeared therefore that the F-C-D/F-G-D-type sets created a stronger sense of diatonic scale conformance than did the F-C-G-type.

The results of Experiments 1 and 2 are a significant departure from existing ideas about creation of a sense of diatonic structure. They specifically refute simple notions of distance of probe from context sequences in the circle of fifths. They support the view that the sense of diatonic structure obtained from three notes of a melody may be related to how far the intervals embodied in these notes are representative of diatonicism.

The question arises as to whether this variable of representativeness is more powerful in creating a sense of diatonicism than is one of the most important stylistic indicators of tonality — the use of major and minor triads (e.g., the sets C-E-G and C-E*b*-G). Major and minor triads form the bedrock of Western tonal music (see e.g., Schenker, 1906/1954) and it might be thought that the use of these structures would create a particularly strong sense of diatonic conformance. For example, Cuddy and Badertscher (1987) have shown that triads are highly successful at establishing tonal hierarchies in which the tonic comes out as most stable and out-of-scale notes as least stable. Yet on the diatonic representativeness variable, triads score lower than F-C-D/F-G-D-type sets. The minor triad (e.g., F-D-A) reflects in the circle of fifths on to the major triad (e.g., F-C-A), and the two together can occur six times (three major, three minor) in a diatonic set (i.e., they have a transpositional-inversional diatonic multiplicity of six). As previously noted, the F-C-D/F-G-D-type set has a multiplicity of eight.

Experiment 3

We carried out an experiment to find out whether sense of diatonic conformance would be greater if derived from the F-C-D/F-G-D-type set or from the F-D-A/F-C-A-type set (major/minor triad). In this experiment, instead of just examining the fittingness ratings of out-of-scale probes we included in-scale probes. We wanted to know whether note sequences which led to out-of-scale probes being judged *less* well-fitting would also lead to in-scale notes being judged *more* well-fitting.

Subjects and methods

The subjects were 21 undergraduate students at London University. None was studying music but previous musical training varied from none to Grade VIII (practical). All subjects underwent all conditions. The basic procedure was as in the previous experiments with subjects listening to a series of three-note sequences, each followed by a probe note, and indicating how well or how badly the probe fitted in with the previous three. There were two independent variables — whether the probe was in- or out-of-scale with the preceding three notes, and whether the pitch-class set was of the type F-C-D/F-G-D or F-D-A/F-C-A (major/minor triad). The out-of-scale probes were randomly selected from the pair F# and C#. The in-scale probes were randomly selected from the pair B*b* and E. Each sequence and associated probe was randomly transposed.

There were 16 sequences in each of the four conditions thus created, giving a total of 64 trials in all. As in previous experiments the order of the notes in the three-note sequences was randomised as was the order of the sequence-types themselves.

The sequences were generated and played on a BBC Model B microcomputer with the sound output diverted from the internal speaker through a low-pass filter to a Leak Model 30 Audio Amplifier and an external speaker. The filter was set at 2.5 KHz (48 dB/8ve) thus attenuating the upper harmonics of the square wave produced by the

BBC and producing a more pleasing sound giving a secure sense of pitch. The notes used ranged upwards for two octaves from middle C, and were played at a rate of two per second with a six second pause between each four-note sequence for subjects to make their ratings, on a seven-point scale.

Results

Table 3 gives the mean ratings for each of the four conditions. For both types of sequence the out-of-scale probes were judged significantly less well-fitting than the in-scale probes, $F(1,20) = 32.8$, $p < 0.001$. The difference between the ratings for out-of-scale and in-scale probes was greater for the F-C-D/F-G-D-type sequences than for the triads, $F(1,20) = 6.5$, $p < 0.025$. This was due to the out-of-scale probes associated with the F-C-D/F-G-D sequences being given lower ratings then those associated with the F-D-A/F-C-A sequences, $F(1,20) = 6.3$, $p < 0.025$. There was no significant difference between the ratings of the in-scale probes associated with the two types of sequence.

Table 3. — *Mean Ratings of Degree of Fit Between In- and Out-of-scale Probe Following Three-note Sequences of the Kind F-G-D/F-C-D Versus F-G-A/F-D-A*

	F-C-D/F-G-D type sequences		F-C-A/F-D-A type sequences	
	In-scale	Out-of-scale	In-scale	Out-of-scale
Mean	4.7	3.5	4.6	3.9
Standard Deviation	0.71	0.82	0.61	0.65

Note. The in-scale probes were judged significantly better fitting than the out-of-scale probes, $F(1,20) = 32.8$, $p < .001$. The difference between the in- and out-of-scale probes was greater for the F-C-D/F-G-D type sequences than the F-C-A-/F-D-A type sequences, $F(1,20) = 6.5$, $p < .025$.

The results showed that as predicted by the representativeness theory, the F-C-D/F-G-D sequences made out-of-scale probes appear less well-fitting even than major/minor triads. However, in-scale probes were judged equally well-fitting in all cases. This suggests an asymmetry in the concept of diatonic conformance, with the fittingness of in-scale notes being subject to different influences from the fittingness of out-of-scale notes.

EXPERIMENT 4

The results of the previous experiments indicated that pitch-class sets of the F-C-D/F-G-D-type created a stronger sense of diatonic conformance (in the sense of enabling out-of-scale notes to be identified) than those of the type F-C-G or the major/minor triad. These results support the view that diatonic representativeness is influential in determining the sense of diatonicism obtained from three-note sequences. In confirmation of this we carried out an experiment which differed from its predecessors in two important respects. Firstly, instead of merely having a comparison between *two* different set-types we compared *three* set-types and specific prediction about the order in which they would be more or less effective in creating a sense of diatonic conformance. Secondly, we modified the subjects' task so that instead of it being merely a question of rating the degrees of fittingness of probe notes, they actually had to identify out-of-scale notes. The purpose of this was to confirm that the pattern of results obtained using the fittingness measure specifically reflected judgments concerning diatonic conformance.

Subjects and methods

Three male and seven female subjects participated in the experiment. None had received formal musical training beyond elementary level although some still played an instrument occasionally for pleasure.

The basic experimental task involved playing pairs of four-note melodies to subjects. Every pair consisted of a three-note sequence followed by an in-scale note and the same three-note sequence followed by an out-of-scale note. After each pair subjects were asked which sequence they thought had contained the out-of-scale note. They made their responses on a seven-point scale as follows: 1 = definitely the first sequence, 2 = almost certainly the first sequence, 3 = probably the first sequence, 4 = don't know, 5 = probably the second sequence, 6 = almost certainly the second sequence, 7 = definitely the second sequence. Before the experiment the concept of "in-scale" and "out-of-scale" notes was described and demonstrated on a piano keyboard. All subjects were able to understand that having played all the white notes within an octave span any of the black notes would be "out-of-scale" in the sense in which we were using the term.

There were three conditions. In the first condition, the three-note set was of the F-C-D/F-G-D-type. In the second condition it was the major/minor triad set. In the third condition it was the F-C-G set.

The out-of-scale note in the sequences prior to transposition (see below) was always F#. The in-scale note was either E or B*b*. Whether the out-of-scale note occurred in the first or second sequence was balanced across the trials. There were four trials in each condition. The order of the trials was randomised and the sequences were transposed at random.

The notes were played using a step-time sequencer linked via a MIDI interface to a Prophet 2000 digital sampling keyboard, outputting to an audio amplifier and speakers. The sound used was a compositely-sampled grand piano (4 samples), sampled at 41 KHz with 12 bit resolution, in a two octave range upwards from middle C. The notes were played at a rate of two per second, with a three-second pause between the sequences in each pair and a six second pause between each pair of sequences for subjects to make their responses.

In trials where the out-of-scale note was in the first sequence of the pair, we subtracted the subjects' ratings from 8. The means were then calculated of the adjusted ratings for each subject in each of the three conditions. These means represented the accuracy of subjects' identification of the out-of-scale note weighted according to their confidence in their judgment.

We hypothesised that if representativeness of diatonic scale structure was important in determining sense of diatonic conformance, then F-C-D/F-G-D sets should give the highest ratings, followed by the major/minor triad set, then the F-C-G set.

Results

The results were as predicted by the theory of representativeness of diatonic scale structure. There was a monotonic trend in the subjects' scores going from the F-C-D/F-G-D-sets to the major/minor triad sets then the F-C-G sets. In order, the scores were 5.4, 4.8 and 4.2 out of seven. The difference between the conditions was significant, $F(2,18) = 6.1$, $p < 0{,}05$, as was the linear trend $F(1,18) = 11.0$, $p < 0{,}01$.

Discussion

The main finding of the four experiments was that systematic differences could be found between the extent to which out-of-scale notes could be identified following different three-note sets, with sets of the type F-C-D/F-G-D being better than major or minor triads which were in turn better that F-C-G sets. These results support the hypothesis that an important factor in determining how pitch information is used to create a sense of diatonic scale conformance is the extent to which the pitch class sets, and hence the intervals, used are representative of diatonic structure. Representativeness here means the number of different places at which pitch-class sets, or patterns of intervals, can occur within a diatonic scale *without* regard to the order in which pitches (intervals) occur. This appears to play a greater role than factors usually considered central to music in the tonal idiom such as the use of major or minor triads. The findings of these experiments appear to provide a parallel in terms of three-note pitch-class sets to Browne's (1981) idea that intervals which are common within the diatonic set are likely to play a role in pattern-matching activities.

Over the four experiments a great many different random orderings of notes within each three-note set type were used, and it is highly improbable that our results were due to factors associated with specific orderings of pitches or intervals. This is not to say that sequential factors are not important in judgments of fittingness in general. For example, as mentioned above, both Deutsch (1984) and Bharucha (1984), as well as Brown (1988), point up the importance of temporal, sequential factors in the abstraction of a tonal hierarchy from an event hierarchy.

Our results suggest the existence, presumably at a non-conscious level, of mental operations on highly abstract representations of pitch relations. More specifically, they suggest that listeners have access to a representation of pitch structure that is somewhat different from anything that has been proposed in the literature so far. It is a representation in which a number of different three-note sets emerge as equivalent despite different orderings of notes and different actual interval values. The most easily-grasped formulation of this representation is in terms of the circle of fifths; as already stated, pitch-class sets are equivalent with respect to diatonic representativeness if they have the same spread within the circle of fifths or are reflections along their mid-points of pitch-class sets with the same spread in the circle of fifths (see Figure 2). However, the most complete formalisations of this representation are those outlined in Forte's (1973) and Balzano's (1980, 1982) descriptions of relations between pitch-class sets, although the factor of the representativeness of a pitch-class set in respect of a superset (here, based on the multiplicity of set-type in respect of the diatonic set) is not explored by either Forte or Balzano, and is only tangentially alluded to by Browne (1981).

It is unclear whether a mental representation of an unordered interval set could exist for *more* than three notes, given the proliferation of possible intervals between four or more notes. Staying with the three-note sets, however, the above line of reasoning would suggest that there may be subjective similarity between equivalent three-note sets in the sense described, although the results of Stammers (1991) would suggest that this is only feasible within the context of such a highly "overlearned" superset as the diatonic set.

Indeed, an important assumption underlying the experiments described here — taking its rationale from the results of Cross, Howell and West (1983) — is the primacy of the diatonic set or diatonic scale. Other "scales" identified as being in common use, such as the harmonic minor, may be considered as deviations from this basic form (see Cross, West & Howell, 1991). This assumption appears to be borne out, at least with three-note sets, by the finding in Experiment 3 that with F-C-A/F-D-A-type sequences our (diatonically) out-of-scale notes (F# and C#) were judged less well-fitting than the in-scale notes (E and B*b*) even though the C# could certainly be accommodated together with F-D-A within a collection of notes forming a harmonic or ascending melodic minor scale. It may nevertheless be that if a "minor" scale form is more clearly established by virtue of more notes being presented, a sense of scale-conformance (of a sort) may arise in a similar way to that found in respect of diatonic structure.

One feature of the results in Experiment 3 bears special mention. We found that a manipulation which influenced listeners' judgments concerning out-of-scale probes did not appear to influence judgments of in-scale probes. The two were not "reflections" of each other, and this phenomenon may require further exploration.

The implications of the present results for more general ideas about how schemas (see e.g., Bartlett, 1932; Norman & Bobrow, 1976) are activated are more speculative, but a new and potentially important factor has emerged. It is possible that the greater the number of different places at which a stimulus pattern can occur within a schema, the more strongly that schema will be activated, other things being equal. This factor could in principle operate in addition to other factors such as the statistical distribution of the occurrence of the stimulus pattern in the environment. This would be in keeping with the view that schema activation occurs by a process of pattern-matching in which schemas compete for activation by stimulus patterns and those that most distinctively overlap with a stimulus pattern at all levels of abstraction succeed. Thus if a given stimulus matches at several different places with a schema, it should have an advantage over patterns which match at only one place.

In conclusion, our series of studies indicates that there are clear consistencies in the effectiveness of different pitch combinations when it comes to activating schemas for diatonic structure, with pitch-interval combinations, or pitch-class sets, which occur at more places within the diatonic set being the most informative. These results suggest a new factor in schema activation more generally, which merits further attention.

References

Balzano, G. (1980). The group-theoretic representation of 12-fold and microtonal pitch systems. *Computer Music Journal, 4*, 66-84.

Balzano, G. (1982). The pitch set as a level of description for studying musical perception. In M. Clynes (Ed.), *Music, mind, and brain: The neurology of music* (pp. 321-352). New York: Plenum.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Bharucha, J.J. (1984). Event hierarchies, tonal hierarchies and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General, 113*, 421-425.

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception, 5*, 219-250.

Brown, H. & Butler, D. (1981). Diatonic trichords as minimal cue-cells. *In Theory Only, 5*, 39-55.

Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only, 5*, 3-21.

Cross, I., Howell, P., & West, R. (1983). Preferences for scale structure in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 444-460.

Cross, I., Howell, P., & West, R. (1985). Structural relationships in the perception of musical pitch. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 121-142). London: Academic Press.

Cross, I., West, R., & Howell, P. (1991). Cognitive correlates of tonality. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure* (pp. 201-243). London: Academic Press.

Cuddy, L.L., & Badertscher, B. (1987). Recovery of the tonal hierarchy. *Perception & Psychophysics, 41*, 609-620.

Deutsch, D. (1982). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (pp. 271-316). London: Academic Press.

Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharucha and Krumhansl. *Journal of Experimental Psychology: General, 113*, 413-416.

Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341-354.

Forte, A. (1973). *The structure of atonal music*. New Haven, NJ: Yale University Press.

Howell, P. West, R., & Cross, I. (1984). The detection of notes incompatible with scalar structure. *Journal of the Acoustical Society of America*, *76*, 1682-1689.

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

Krumhansl, C. L. (1990). *Cognitive foundations of tonal harmony*. Oxford: Oxford University Press.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 579-594.

Norman, D., & Bobrow, D. (1976). On the role of active memory processes in perception and cognition. In C. Cofer (Ed.), *The structure of human memory* (pp. 114-132). San Francisco: Freeman.

Schenker, H. (1906). *Harmony*. Chicago: University of Chicago Press. (E.M. Borgese, Trans., 1954)

Simonton, D. (1984). Melodic structure and note-transition probabilities: A context analysis of 15,618 classical themes. *Psychology of Music*, *12*, 3-16.

Stammers, D. (1991). *The cognitive representation of set-theoretic structure in musical pitch relations*. Unpublished M.Ph. thesis, University of Cambridge, Cambridge.

Egham Hill, Egham, Surrey
United Kingdom

Cambridge
United Kingdom

Gower Street, London
United Kingdom