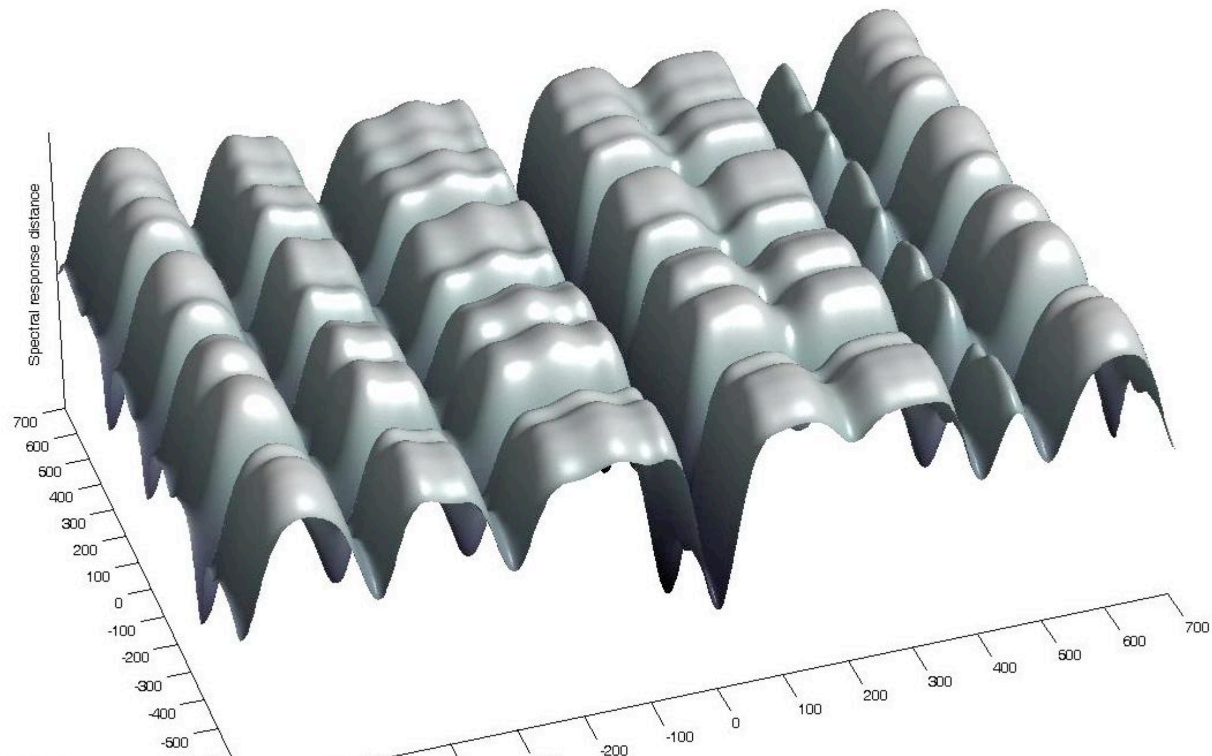
A Psychoacoustic Model of Harmonic Cadences
Pro Gradu Thesis

Music Mind Technology
(MA Programme, University of Jyväskylä)

Andrew J. Milne

2007–2009



JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanities	Laitos – Department Music
Tekijä – Author Andrew J. Milne	
Työn nimi – Title A Psychoacoustic Model of Harmonic Cadences	
Oppiaine – Subject Music, Mind and Technology	Työn laji – Level Master's Degree
Aika – Month and year July, 2009	Sivumäärä – Number of pages 92

Tiivistelmä – Abstract

This thesis presents a psychoacoustically derived computational model of the perceived distance between any two major or minor triads, the degree of activity created by any given pair of triads, and the cadential effectiveness of three-triad progressions. The model is tested against conventional music theory, and ratings given by thirty-five participants for the "similarity" and "fit" of triads in a pair, and the "cadential effectiveness" of three-triad progressions. Multiple regressions show that the model provides highly significant predictions of the experimentally obtained ratings. Finally, it is argued that because the model is based upon psychoacoustic axioms, it is likely the regression equations represent true causal models. As such, the computational model and its associated theory question the plausibility of theoretical approaches to tonality that use only long-term memory and statistical features, as well as those approaches based upon symmetrical geometrical structures like the torus. It is hoped that the approach proposed here may herald not only the return of psychoacoustics to tonal music theory, but also the exploration of the tonal possibilities offered by non-standard tunings and non-harmonic timbres.

Asiasanat – Keywords music cognition, tonality, harmony, cadences, psychoacoustics, voice-leading

Säilytyspaikka – Depository

Muita tietoja – Additional information

ABSTRACT

This thesis presents a psychoacoustically derived computational model of the perceived distance between any two major or minor triads, the degree of activity created by any given pair of triads, and the cadential effectiveness of three-triad progressions. The model is tested against conventional music theory, and ratings given by thirty-five participants for the "similarity" and "fit" of triads in a pair, and the "cadential effectiveness" of three-triad progressions. Multiple regressions show that the model provides highly significant predictions of the experimentally obtained ratings. Finally, it is argued that because the model is based upon psychoacoustic axioms, it is likely the regression equations represent true causal models. As such, the computational model and its associated theory question the plausibility of theoretical approaches to tonality that use only long-term memory and statistical features, as well as those approaches based upon symmetrical geometrical structures like the torus. It is hoped that the approach proposed here may herald not only the return of psychoacoustics to tonal music theory, but also the exploration of the tonal possibilities offered by non-standard tunings and non-harmonic timbres.

Acknowledgments

Bill Sethares for being such an available and patient sounding board, for profound discussions about musical perception and mathematics, and in particular for supplying me with the MATLAB routines that served as the starting point for my model.

Tuomas Eerola for welcoming me with open arms into academia, and allowing me to follow my own path for my Master's thesis when a safer choice would have been much, mmm, safer for all of us. Also to Petri Toiviainen for continuing this policy of openness, and in particular for suggesting the use of the cosine distance. And to Vinoo Alluri for lighting up the darker corners of MATLAB's functionality. MMT rules!

Jim Plamondon for extensive proof reading, but most of all for being Jim: making the contacts and introductions that made this project possible, and for his indefatigable enthusiasm and optimism, and belief in me.

To Ken for helping to fund my studies, and to Christine and Astra for graciously giving me the time to follow my own private obsession.

TABLE OF CONTENTS

LIST OF TABLES vii

LIST OF FIGURES viii

1. INTRODUCTION 1

2. BACKGROUND 3
   - 2.1. Tonality and Cadences 3
   - 2.2. Theories of Harmonic Tonality 3
     - 2.2.1. The Generative Tonic 4
     - 2.2.2. The Central Tonic—Structuralism 7
     - 2.2.3. The Close Neighbour Tonic—Voice-Leading 9
     - 2.2.4. The Resolved Tonic—Melodic Activity or "Dissonance" 10
     - 2.2.5. The Familiar Tonic—Long-Term Memory 11
     - 2.2.6. The Dialectical Tonic—Functionalism 13
   - 2.3. An Ideal Theory of Harmonic Tonality 15

3. THE PSYCHOACOUSTIC MODEL 16
   - 3.1. The Model's Psychoacoustic Variables 18
     - 3.1.1. Pitch Distance (*pd*) 18
     - 3.1.2. Fundamental and Spectral Response Distances (*frd* and *srd*) 18
       - 3.1.2.1. Spectral Response Distance 20
       - 3.1.2.2. Fundamental Response Distance 21
   - 3.2. The Model's Cognitive Variables 22
     - 3.2.1. Voice-Leading Distance (*vld*) 24
     - 3.2.2. Spectral Distance (*sd*) 25
     - 3.2.3. Tonal Activity (*act*) 27
       - 3.2.3.1. Continuous Activity (act_c) 28
       - 3.2.3.2. Discrete Activity (act_d) 29
       - 3.2.3.3. Asymmetry of Tonal Activity 30
       - 3.2.3.4. Exclusive and Non-Exclusive (Double) Alterations 31
   - 3.3. Applying the Model to Harmonic Cadences 32
     - 3.3.1. Cadential Form—Patterns of Tonal Activity 32
     - 3.3.2. Cadential Fit—Spectral Distances 33
     - 3.3.3. Cadential Salience and Flow—Voice-Leadings Toward Resolution 33
       - 3.3.3.1. Position in Final 33
       - 3.3.3.2. Direction 34
       - 3.3.3.3. Amount 34
     - 3.3.4. Cadential Synergy—Latent Activities of the Embedding Scale 34
     - 3.3.5. Cadential Complexity—Melodic Complexity of Embedding Scale 35
     - 3.3.6. The Model's Calculations 35
   - 3.4. Interpreting the Model 36

3.4.1. Plotting the Model's Calculations 36
3.4.2. Charting Tonal Activities 40

4. NON-EXPERIMENTAL TESTING OF THE MODEL 42
4.1. Cadences 42
4.1.1. IV→V→ Cadences 43
4.1.2. ii→V→ Cadences 44
4.1.3. iv→V→ Cadences 44
4.1.4. ♭II→V→ Cadences 45
4.2. Tonal Functionality 46
4.3. Tonal Asymmetry 46
4.4. Tonal Dualism 47
4.5. Tonal Scales 48
4.6. Tonal Harmony 49
4.7. Tonal Robustness 49
4.8. Discussion and Conclusion 50

5. EXPERIMENTAL TESTING OF THE MODEL 51
5.1. Method 51
5.1.1. Participants 51
5.1.2. Apparatus and Procedure 52
5.1.3. Stimuli 52
5.1.3.1. Triad Pairs—"Similarity" and "Fit" 53
5.1.3.2. Triad Triples—"Cadential Effectiveness" 53
5.2. Results 56
5.2.1. Similarity 56
5.2.2. Fit 58
5.2.3. Similarity and Fit 61
5.2.4. Cadential Effectiveness 63
5.3. Discussion and Conclusion 66

6. DISCUSSION AND CONCLUSION 69

REFERENCES 71

APPENDIX A: NOTATIONAL STYLE 74
Absolute Notation 74
Relative Notation 74
Progressions and Pairings 74

APPENDIX B: MATHEMATICAL PROOFS AND DERIVATIONS 75
Cosine Distance Between Two Response Curves 75
Approximate Spectral Distance Between Two Tones with Harmonic Partials 76

APPENDIX C: TRIAD PAIRS AND TRIPLES 78

APPENDIX D: INTERFACES OF THE TWO EXPERIMENTS 80

APPENDIX E: INTER-PARTICIPANT CORRELATION MATRICES 82

LIST OF TABLES

Table 5.1. Pearson correlations, and their one-tailed significance, between *sim*, *bas*, *ten*, *alt*, *sop*, *frd*, $act_c(X \leftrightarrow Y)$, and $act_d(X \leftrightarrow Y)$. 57

Table 5.2. Regression coefficients and significance for multiple regression of *sim* on *frd*, *bas*, and $act_c(X \leftrightarrow Y)$. 58

Table 5.3. Pearson correlations, and their one-tailed significance, between *fit*, *srd*, $act_c(X \leftrightarrow Y)$, and $act_d(X \leftrightarrow Y)$. 59

Table 5.4. Regression coefficients and significance for multiple regression of *fit* on *srd* and $act_c(X \leftrightarrow Y)$. 60

Table 5.5. Pearson correlations between *sim*, *fit*, *frd*, and *srd*. 61

Table 5.6. Example triad pairs are shown in rank order, from smallest to largest, of their *vld* (as calculated with Eq. (5.1)). Common neo-Riemannian abbreviations representing some of the transforms are in the second row. 62

Table 5.7. Pearson correlations, and their one-tailed significance, between *eff*, $act_d(A \rightarrow P|\mathbf{P})$, $act_d(A \leftarrow P|\mathbf{P})$, $act_d(P \rightarrow F|\mathbf{P})$, $act_d(P \leftarrow F|\mathbf{P})$, $act_d(A \leftarrow F|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$. 64

Table 5.8. Regression coefficients and significance for multiple regression of *eff* on $act_d(A \rightarrow P|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$. 65

Table C.1. The triad pairs used in the experiment to get ratings for "similarity" and "fit". The pitch of each pair was randomised over 12 equally tempered semitones. The spectral response distances (*srd*) calculated by the model are also shown. 78

Table C.2. The triad triples used in the experiment to collect ratings for "cadential effectiveness". The pitch of every progression was randomised over 12 equally tempered semitones. The Group number (see Sect. 5.1.3.2) is also shown. 79

Table E.1. Inter-participant correlations (Pearson) for ratings of "similarity". The single outlying participant is not shown. The column to the right shows the per participant mean correlation, the value at the bottom right shows the overall mean correlation. 82

Table E.2. Inter-participant correlations (Pearson) for ratings of "fit". The three outlying participants are not shown. The column to the right shows the per participant mean correlation, the value at the bottom right shows the overall mean correlation. 83

Table E.3. Inter-participant correlations (Pearson) for ratings of "cadential effectiveness". The column to the right shows the per participant mean correlation, the value at the bottom right shows the overall mean correlation. 84

List of Figures

Figure 3.1. A path diagram showing the proposed flow of causation from the psychoacoustic variables pitch distance (*pd*), fundamental response distance (*frd*), and spectral response distance (*srd*), to the cognitive variables of voice-leading distance (*vld*), spectral distance (*sd*), and tonal activity (*act*). Error terms are not shown. 16

Figure 3.2. The spectral response distances between a 12-TET reference triad and a continuum of differently tuned triads, plotted against their pitch distances. The reference triad is major in (a), (b), and (c), minor in (d), (e), and (f). In (a) and (d), the horizontal axes show the pitch distance (in cents) between the "roots" and "fifths" of the continuum triads and the root and fifth of the reference triad; the vertical axis shows the pitch distance between the "thirds" of the con tinuum triads and the reference triad. The greyscale indicates the spectral response distance (the lighter the colour the greater the spectral response distance). The remaining figures view (a) and (d) from the side, so their vertical axes show the spectral response distance; in (b) and (e) the horizontal axes are equivalent to the horizontal axes of (a) and (d); in (c) and (f), the horizontal axes are equivalent to the vertical axes of (a) and (d). A selection of continuum triads is labelled. 38

Figure 3.3. A chart of the tonal activities of all 12-TET triads in reference to a C or c triad. The vertical arrows represent activities due to **P** comparisons. Prototypical continuum triads are circled. 40

Figure 5.1. Multiple regression of *sim* on *bas*, *frd*, and $act_c(X \leftrightarrow Y)$. 58

Figure 5.2. Multiple regression of *fit* on *srd* and $act_c(X \leftrightarrow Y)$. 60

Figure 5.3. A path diagram showing the proposed relationships between the cognitive variables discussed above—including a long-term memory (*ltm*) component—and the measured variables similarity (*sim*) and fit (*fit*). Error terms are not shown. 62

Figure 5.4. Multiple regression of the cadential effectiveness (*eff*) of 72 different triad triples on $act_d(A \rightarrow P|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$. 65

Figure D.1. GUI of the first experiment. 80

Figure D.2. GUI of the second experiment. 81

1. INTRODUCTION

Psychoacoustic approaches have provided relatively effective explanations for why certain simultaneities of notes (chords) are typically considered "dissonant" while others are considered "consonant" (notably the major and minor triads that are so important in both the theory and practice of Western tonal music) (Helmholtz, 1877; Plomp & Levelt, 1965; Kameoka & Kuriyagawa, 1969; Parncutt, 1988; Sethares, 2004). However, to date, there has been no psychoacoustic explanation for one of the most important and mysterious aspects of Western tonal music—the fact that a succession of consonant chords can induce feelings of "expectation" and "resolution" that are not produced when the same chords are played in isolation, or in a different order.

For example, listeners will typically feel that in the chord progression F major→G major→C major, the second chord sounds particularly expectant whereas the third chord resolves this expectation, thus providing a sense of closure. Chord progressions such as these are called cadences, and they are typically used in tonal music to mark the ends of phrases, or entire sections. Interestingly, cadences are commonly constructed with only consonant triads (the example above is the familiar IV→V→I cadence; other common cadences using only major and minor triads are ii→V→I, iv→V→i, and iv→V→I). Such cadences imply that the expectation or resolution induced by a chord is not necessarily a function of its inherent (vertical) consonance or dissonance, but rather of its temporal (horizontal) context—most particularly the chords that directly precede and proceed it.

Any theory of *harmonic tonality*—that form of music using chords (principally triads) to establish a tonic ("home") note or triad (Krumhansl, 1990)—must provide an explanation for these feelings of expectation and resolution that lie at its very heart. In the absence of successful psychoacoustic theories to account for this phenomenon, many contemporary researchers have suggested a statistical (long-term memory) explanation (Bharucha, 1987; Krumhansl, 1990; Tillmann, Bharucha, & Bigand, 2000; Levitin, 2006). These approaches suggest that we are culturally trained, by exposure, to expect certain progressions, and this accounts for the effect produced by the regularities (such as cadences) that are found in tonal music—that is, if we've heard it before, we expect to hear it again. There is little doubt that this is a credible approach, but it has a number of problems if used as the sole explanation for these effects. For example, (a) it implies that the effect induced by a given chord progression—such as a cadence—should be very plastic, but there is little evidence, from either a cultural or historical perspective, that this is the case; (b) short-term memory has been demonstrated to play a significant role in perception of tonality (Leman, 2000); (c) typical cadential progressions have been readily adopted, with no modification, by non-Western cultures (e.g., see Agawu (2003)).

Statistical approaches undoubtedly play an important part in the cognition of harmonic cadences, but I propose there are important psychoacoustic processes that underlie them. In this thesis I present a psychoacoustically derived model designed to explain the flow of expectation and resolution induced by a succession of triads (Sect. 3). The model is built in MATLAB, and is currently relatively simple (it calculates only root-position major and minor triads), and can be substantially developed. I compare the model's predictions to a broad range of regularities identified by conventional tonal music theory (Sect. 4), and to experimentally obtained human ratings of the "similarity", "fit", and "cadential effectiveness" of a variety of triad progressions (Sect. 5). Both data sets provide strong support for the model. I finish with a discussion of some of the implications of this proposed psychoacoustic approach to tonal music theory and practice (Sect. 6). The following section (Sect. 2) provides the context and background for the model and its associated theory by giving a thematic summary of existing theories of harmonic tonality and cadences.

But, before proceeding, let me give a quick explanation of the notation used in this thesis: for brevity, I will refer to major triads in upper case, minor triads in lower case—so "A" is an A major triad, "g" is a g minor triad. Furthermore, without qualification, all triads are considered to be major or minor and in root position. Notes are differentiated from chords by being written in italic (lower case)—so "*a*" and "*g*" are notes, not triads. For an explanation of the uses of Roman numerals (for triads) and Arabic numerals for scale degrees, please see Appendix A.

2. BACKGROUND

2.1. Tonality and Cadences

Tonality is a term with many overlapping meanings, but its most common usage is reasonably consistent (Dahlhaus, 1980; Hyer, 2002): "the term denotes, in the broadest sense, relationships between pitches, and more specifically between pitches having a 'tonic' or central pitch as its most important element" (Dahlhaus, 1980, p. 52). Dahlhaus also gives a concise definition for the compound term *harmonic tonality*: it refers to a particular form of tonality that is "determined by chordal relationships, which formed the foundation of composition from the 17th century to the early 20th century" (1980, p. 53). Similarly, Krumhansl defines tonal-harmonic music as "traditional Western music…that is tonal in the sense of being organized around a central reference pitch (tonic) and…harmony is important for establishing the tonal framework" (1990, p. 9).

The harmonic cadence is typically considered to be one of the defining characteristics, even the source, of harmonic tonality. For example, Lowinsky (1961) wrote that "the cadence is the cradle of tonality" (p. 4); and harmonic cadences, such as IV→V→I, ii→V→I, and iv→V→i, and so forth, do indeed exemplify, in the most distilled form, those patterns of expectation and resolution that occur in the harmonic tonality. It seems likely, therefore, that understanding the cadence—how it is that one chord out of a progression becomes *tonicised* (i.e., given a sense of repose)—is a key step towards understanding tonality as a whole.

2.2. Theories of Harmonic Tonality

The quest for those (hopefully simple) principles that underlie the regularities (notably harmonic cadences) of tonal-harmonic music, has been ongoing since the birth of tonality itself, and is still an area of vigorous debate and research. An analysis of the theoretical approaches taken by music theorists (both historical and modern) to answer such questions shows that only a relatively small number of different underlying principles have been used (though some theorists use a combination of them). In this section, I briefly go through each of these principles; indicate their origin, and the arguments used to support them. I also include critical analysis of each of the principles, indicating their shortcomings. I conclude with a description of some of the features an ideal theory of tonality should have.

(Please note that I do not consider the impact of metrical position upon the effectiveness of a cadence—cadences where the tonic falls on a strong beat are typically considered stronger than those ending on a weak beat—because this is an area outside the scope of this thesis.)

2.2.1. The Generative Tonic

The principle of a generative tonic is frequently encountered in music theory—forms of it are found in theorists as diverse as Rameau, Schenker, and Mathieu. The generative tonic theory asserts that the harmonic partials of a single tone (the tonic) "generate" other structurally important tones, chords, or even keys, of a musical system. When these generated tones, chords, or keys, return to the tonic they are, therefore, returning to their "source". This is theoretically enticing because it has both a clear acoustic basis (the harmonic series), and a clear psychological basis (the return to the source explains the feeling of "finality", or "completion" induced by the tonic). But the theory has significant problems when followed through to its natural conclusion. These issues are explored below.

In the 18th century, Rameau used the generative principle to explain the tonic's different levels of repose depending on whether it is approached from its dominant (a "perfect cadence") or from its subdominant (an "irregular cadence") (Caplin, 1983). The perfect cadence is stronger because (unlike the irregular cadence) the fundamental bass (i.e., root) of the triad "returns to its source" (Rameau quoted in Caplin (1983, p. 4)). Rameau's use of the generative tonic principle seems to produce a reasonable and uncontroversial conclusion—the authentic cadence is stronger than the plagal. But if downward fifth progressions provide a sense of conclusion on the second chord, when does this process end; one could resolve G to C, which then resolves to F, which resolves to B♭, which resolves to E♭, which resolves ..., ad infinitum? The only way to stop this infinite cycle of resolutions is to make some sort of scalic restriction; for example, we might say that all chords can come only from a single diatonic scale—this would mean that G resolves to C, which resolves to F, which cannot resolve to B♭ because B♭ is not in the same diatonic scale as G, C, and F. This implies that, in this context, F is the tonic; but this contradicts conventional music theory, which would give C as the tonic. So, in this vital respect, the theory of the generating tonic fails to provide an accurate prediction (see Lester (2002, p. 771), for a further discussion of the theoretical difficulties posed by Rameau's *corps sonore*).

Schenker used the generative principle to explain many aspects of music—chords, scales, and key relationships—though voice leading concerns became more important in his later writings. He claimed (1954) that "tonicalization can be effected only by a process of inversion—tonicalization being essentially a descent to the tonic!" (p. 261). By "inversion" and "descent" he is referring to a chord progression in which the root tone of a chord takes a more "humbling" position as third or fifth in the subsequent chord (p. 235). This suggests the following progressions should be cadential: V→I, V→i, v→I, v→i, iii→I, III→I, ♭III→i, ♭iii→i. In conventional music theory and practice, however, only the first two of these progressions is typically treated as cadentially effective (and furthermore the V→I progression is only cadential in the presence of the subdominant, otherwise its true nature is ambiguous—it could be I→IV).

Schenker also used a scalic explanation for tonality. He wrote that any given tonic corresponds to (it generates) a unique diatonic scale (the major, or Ionian, mode); a logical consequence of this is that any given diatonic scale must have a unique tonic chord. However, his explanation for why a specific mode of the diatonic scale is generated by the tonic—that mode produced by going down one fifth and up five fifths from the tonic scale degree (i.e., the Ionian mode)—is weak. To explain the five ascending fifths, Schenker (1954) used an explanation based on an unjustified (and somewhat dubious) perceptual premise:

> If he [the artist] did not want to lose sight of his point of departure, he had to restrict himself to the use of only five tones above the *C*. Here, again, human perception has wonderfully respected the limit imposed by the number five. (p. 30)

To explain the single descending fifth, Schenker (1954) used nothing more than a metaphor that, even at a metaphorical level, seems unconvincing:

> The inversion of the fifths, leading to a fivefold descent down to the tonic, entailed a new consequence: The artist, face to face with the tonic, felt an urge to apply inversion once more, searching, so to speak, for the ancestor of this tonic with its stately retinue of tones. Thus he discovered the subdominant fifth F, which represents, metaphorically speaking, a piece of past history of the tonic C. (p. 38)

But why not, for example, include the *b*♭ ancestor of the ancestral *f*, why stop at the "parent" rather than the "grandparent"?

Mathieu's theory assumes a similar generating tonic (though Mathieu supposes that 7- and higher-limit harmonics are important). He also requires similar metaphysical explanations. Reviewing Mathieu's Harmonic Experience, Carey (2002) writes:

> Tonal systems entail relationships above and below the tonic, ("overtonal" and "reciprocal" are Mathieu's terms), but like a run of Biblical "begats," the overtones run asymmetrically, in one direction only. Where do the reciprocals come from? Propagation metaphors falter. If the tonic is also the generator, metaphysics must be summoned to explain the riddle of the subdominant. "When you sing F you create C. How can you create the creative principle? How does one go about giving birth to a musical god? That is the work of the Musical Mother [the syllable ma stands for the fourth degree of the scale] ... You who dare to sing F in the C world become the embodiment of the creative and the sacred". (p. 123)

The fundamental problem (pun intended) with the principle of the generating tonic is the subdominant. The theory would have the subdominant as a tonic (unless metaphysical wriggles, such as those above, are applied), and would therefore suggest that a subdominant chord anywhere in the vicinity of a V→I cadence would weaken that cadence (because it would remind the listener that there is a "better" resolution chord). But in actual musical practice, not only does this not happen, the subdominant actually plays an important role as a preparation to the dominant in harmonic cadences—indeed it actually seems to strengthen the cadential effect. Schenker himself, writing about that paragon of tonal practitioners, J. S. Bach, states, "it was almost a rule for him to anchor his tonic, right at the outset, by quoting, first of all, the subdominant and then the dominant fifth, and only then to proceed with his exposition" (1954, p. 38); and Agmon (1996) argues that:

> Even in terms of Schenker's own theory, the idea that IV or II^6 represent some sort of "leaping passing-tone" configuration in the bass surely leaves much to be desired. ... I have often heard the complaint that a musical phrase is robbed of its essence once the "structural subdominant" is removed. (para. 21)

2.2.2. The Central Tonic—Structuralism

The use of regular geometrical structures to represent the interrelationships, and "distances", between different notes, chords, or keys has a long history, including Euler's *Tonnetz* (Euler, 1739), Weber's table of the relationships of keys (see Bernstein (2002)), Hauptmann's conception of the major key (see Harrison (1994)), Oettingen's tonal space (see Klumpenhouwer (2002)), Schoenberg's chart of the regions (1969), Krumhansl's spatial representations of interkey distance (1990), Chew's spiral array (2000), Lerdahl's tonal pitch space (2001), and so on. These structures are generally lattice-like, in that they have translational symmetry, and can (in principle) have infinite extension.

The concept of the tonic being the centre of such a structure is an important theoretical strain within such structural approaches, but a geometrical structure can only have a centre if it is truncated in some way (i.e., it has finite extent). A typical truncated structure is the chain of six fifths, which can be used to generate the seven notes of the diatonic scale. The two most central triads in this chain are I and vi—a conclusion that chimes well with conventional musical theory and practice. For example, in the chain of fifths *f*, *c*, *g*, *d*, *a*, *e*, *b*, the two most centrally located triads are C (*c*-*e*-*g*) and a (*a*-*c*-*e*).

Like the theory of the generating tonic, the theory of the central tonic provides an elegant (though different) psychological explanation—the tonic is the centre of the system, it is equally balanced between being a generator and having been generated. The tonic, therefore, represents the most balanced, central, reference point in the system, and in the cadence, the music proceeds from this centre to one extreme (the subdominant), skipping over the centre to the other extreme (the dominant), with a final return to the point of balance (the tonic). Riemann wrote, "it is a fact that successions of chords further related from one another stimulate the expectation of a mediating chord which is closer related to both chords" (Mickelsen, 1977, p. 37).

From a cognitive point of view, this theory requires that this generated chain of fifths is somehow intelligible to the mind, intelligible enough that we "feel" the natural balance of the centre. Whether this is actually possible, however, is questionable—for example, Dahlhaus (1990) writes:

> One can think at the same time of the fourth and fifth in reference to the whole tone, or of the doubling of the whole tone in reference to the major third, but not of four fifths in reference to the major third. (pp. 167–168)

Beyond the question mark that hangs over of the intelligibility of the complete chain of fifths, there are other serious issues with the principle of centrality:

1. The actual centre of the diatonic chain of fifths is scale degree 2 (i.e., in the chain extending from *f* to *b*, the central note is *d*)—the chords I and vi are both slightly, and equally, off-centre. Despite this, scale degree 2 does not seem to have any sense of tonicness, while I and vi do function effectively as tonics.

2. For some non-diatonic scales, the centrality test fails to predict sensible tonics. For example, the central triads of the ascending melodic minor scale (e.g., *e*♭, (*b*♭), *f*, *c*, *g*, *d*, *a*, (*e*), *b*) are F and d, rather than the expected c (or possibly G).

3. The I and vi chords are equally distant from the absolute centre, but the former is typically considered a much stronger tonic than the latter.

4. Similarly, the iii chord has a similar location in the chain of fifths to the V chord, but does not seem to function effectively as a cadential penult (e.g., ii→V→I and IV→V→I cadences are much more common than ii→iii→I and IV→iii→I cadences).

5. Because these geometrical representations are symmetrical, they cannot explain the temporal asymmetries of cadential chord progressions (and other aspects of tonality). For example, in the chain of fifths, the IV and V chords are equidistant from I—this fails to explain how it is that the V is generally felt to be more "expectant" than the IV, and that the majority of cadences follow a IV→V→I pattern, not a V→IV→I pattern.

Such temporal asymmetries (which have been demonstrated in numerous experiments, e.g., Brown (1988), Cuddy & Thomson (1991), Toiviainen and Krumhansl (2003)) are an important part of human perception of tonality, and the inability of a symmetrical structure (of however many dimensions) to account for these asymmetries is well acknowledged (see, e.g., Krumhansl (1990) and Woolhouse (2007)).

2.2.3. The Close Neighbour Tonic—Voice-Leading

Voice-leading principles are a core part not just of Schenkerian approaches, but also of Lerdahl's tonal pitch space. The voice-leading principle asserts that the "tonicness" of a putative tonic can be strengthened (or weakened) by the size of the interval (or intervals) by which it is approached. For example, Schenker gave the following voice-leading explanation for the tonicising effect of the V→I progression:

> There is no doubt that it has been the purely contrapuntal collaboration of leading tone principles together with the demand for completeness of triads that has introduced us for the first time to the dominant-concept in any form. (1987, p. 47)

Schenker related the strength of the voice-leading effect to the size of the steps involved, noting that that the major dominant has a semitone leading tone, while the minor dominant (which is usually avoided in cadences) does not.

More generally, this suggests that in a scale such as the diatonic, (e.g., *c*, *d*, *e*, *f*, *g*, *a*, *b*), those notes that can be approached by semitone (i.e., *b*, *c*, *e*, and *f*) are likely to be tonics, while those notes that are approached only by whole tones (i.e., *d*, *g*, and *a*) are not likely to be tonics. However, the voice-leading principle lacks *directionality*—it gives no inherent guidance as to whether *b* leads to *c* or *c* leads to *b* (or if *e* leads to *f*, or *f* leads to *e*); for example, if *c* leads to *b* and *f* leads to *e*, this suggests that iii can function as an effective tonic (but this is contradicted by conventional music theory and practice). For this reason, the voice-leading principle can be only a part of a broader theory. For example, it requires another factor to suggest a putative tonic, which is confirmed (or not) by the voice-leading; or a factor that gives a preferred direction to the voice-leading.

This is precisely the approach taken by Lerdahl (2001)—his melodic and harmonic attraction rules say that the attraction between any two successive chords (or tones) is inversely proportional to their semitone distance (or squared distance), respectively, and the direction of attraction proceeds from a tone with a lower *anchoring strength* to a tone with higher anchoring strength (2001, pp. 161–162).

Furthermore, although voice-leading principles give a neat explanation for why the IV→V→I cadence is favoured over the V→IV→I and IV→v→I cadences, they fail to explain why

♭vii→I (the Phrygian cadence) is not considered as cadentially effective as V→I (in both V→I and ♭vii→I, the tonic scale degree is approached by one tone and one semitone). Neither do they explain the rarity of the ♭II→I cadence despite it having three semitone-sized leading tones.

2.2.4. The Resolved Tonic—Melodic Activity or "Dissonance"

Rameau believed the resolution from harmonic dissonance to harmonic consonance—as exemplified by the V^7→I progression, where the dissonant third and seventh of the first chord resolve to the root and third of the second chord—provided an explanation for the sense of repose induced by the tonic triad.

This theory seems quite reasonable, but is limited in its applicability because it only covers the use of a harmonically expressed tritone (i.e., the 4 and 7 are expressed simultaneously), and so cannot directly explain cadences that use only major and minor triads (like ii→V→I and IV→V→I).

Two centuries later, Hindemith hinted at an expansion of this concept:

> The root-succession in which the tonal center is preceded by its fourth and its fifth forms the ideal cadence. What makes it ideal is not only the succession of closely related tones. For the chords built on the fourth and the fifth embody (at least when they are simple triads, the one on the fifth being major) a tritone divided between them, which is resolved in the final chord. … This same tritone relation results from the cadential root-progression major second—fifth—tonic, and accordingly this root progression is also very strong. (1942, pp. 139–140)

Neither of the tritones in the above examples are harmonically expressed—each is a melodic tritone played in different voices of two successive chords. It is clear that Hindemith believed this resolution of a melodic tritone is associated with the cadential effectiveness of the IV→V→I, or ii→V→I, progressions; though it is less clear if he also believes the tritone's resolution actually is a cause of cadential effectiveness.

A principle that the resolution of either melodic or harmonic tritones is the source of tonicity seems to carry a lot of explanatory power. Most conventional cadential patterns (even uncommon ones) can be explained by the resolution of a melodic tritone, and most resolutions of the tritone do actually produce effective cadences. Such a theory requires the introduction of some sort of function of *melodic "dissonance"* or *activity* (i.e., that certain intervals, whose tones are expressed successively, are inherently unstable in some way). An important advantage of such a melodic dissonance function is that it might be possible to psychoacoustically derive the dissonance of melodic dyads, analogous to how the dissonance of harmonic dyads can be derived from a sensory dissonance function (e.g., Plomp & Levelt (1965)).

But the concept of melodic dissonance also has problems—it does not explain why melodic versions of some harmonic dissonances (e.g., tritones) are active and require resolution, while others (e.g., diatonic semitones) do not (e.g., the melodic minor second $b \rightarrow c$, between the chords G and C, does not require resolution to $c\sharp$).

2.2.5. The Familiar Tonic—Long-Term Memory

Long-term memory (also known as statistical or schematic) approaches are common in contemporary music cognition research, being the explanation favoured by, for example, Krumhansl, Bharucha, and Lerdahl.

Lerdahl justifies the importance of the diatonic scale on the basis of Balzano's uniqueness, coherence, and simplicity, but to explain the privileging of the Ionian mode of the diatonic scale, he relies upon a long-term memory explanation: "exposure to music is a prerequisite for internalizing a tonal hierarchy" (2001, p. 41)—that is, we privilege the Ionian mode because we are familiar with it. Considering that the Ionian mode forms the foundation of his basic tonal pitch space, which in turn determines the anchoring strength of tones, this seems a somewhat insubstantial justification.

Krumhansl's (1990) empirical data show that a statistical approach provides a more effective explanation than at least one psychoacoustic approach:

> Although the acoustic properties associated with consonance may determine to some extent both the way tones are used in music and the quantified tonal hierarchies, the latter two variables are more strongly interdependent. These results point to learning as the dominant mechanism through which the tonal hierarchies are formed. (p. 76)

But many other psychoacoustic approaches are possible than the specific tonal consonance method tested here, and it is wise to remember the dictum that correlation does not imply causation. For example, it is likely that any underlying cause (e.g., psychoacoustic) of tonal hierarchies would also be a direct cause of the note-choices made by composers and musicians; in which case, we would expect a high correlation between the way tones are used in music and tonal hierarchies, even in the absence of any causal relationship between them.

Of his MUSCAT model, Bharucha (1987) writes that its units, and the strength of the links between them, are assumed to depend on learning:

> It would be surprising if chord and key units existed innately. A more plausible hypothesis is that a hierarchical organization of units characterizes a general-purpose cognitive architecture that finds, over time, the mapping of low-level units to high-level units that produces the optimal patterns of expectation. Thus the same neonatal networks exposed to different cultures will develop different unit mappings, but they will all have a hierarchical structure. (p. 24)

However, he also writes:

> The development of musical structures may, in some cultures, have been biased in the direction of a preference for the sensory consonance these relationships engender. Furthermore, the ubiquitous co-occurrence of frequencies in the harmonic series … may cause the strengthening of connections between representations of these frequencies. (1987, p. 27)

Both Krumhansl and Bharucha, therefore, accept that acoustic explanations may play some part in constraining, or biasing, learning towards certain psychoacoustically motivated structures. But they both consider long-term memory to be the dominant factor.

Putting constraints onto such learning is not just a theoretical hedge: it is an epistemological necessity. A theory based purely on long-term memory states that any pattern to which a person is repeatedly exposed will become internalised such that that person will expect similar patterns. This seems reasonable (though somewhat Pavlovian), but it also means that potentially any pattern could be internalised and, therefore, any possible musical system is possible. This raises an important issue—can such a theory be falsifiable? For example, if a musical system is not used anywhere in the world, a purely long-term memory explanation must be that this is because nobody has been exposed to it; if a musical system is used somewhere in the world, that is presumably because people have been exposed to it; these are circular arguments, which mean the theory cannot be tested against existing examples of factual music (or the absence of existing examples of factual music). This is quite different to a psychoacoustic theory whose predictions can be tested against factual musical examples from anywhere in the world.

For this reasons, long-term memory approaches must have some constraints—be they psychoacoustic, neurological (e.g., Bharucha (1987)), or cognitive—which will predict a limited set of possible, or likely, musical systems that can be tested against actual musical practice.

There are three further problems with a purely long-term memory explanation:

1. It would imply that the rules of common practice harmony should be plastic (like the brain), but the evidence is that the basic rules of harmony (e.g., that IV→V→I is a more effective cadence than, for example, V→IV→I, or iii→ii→I) are actually immutable; despite the valiant efforts of generations of composers to be original, they have singularly failed to break or supplant these "old" rules.

2. Western harmonic tonality has been readily accepted, incorporated, and understood by cultures that previously had no contact with it (see, e.g., Agawu (2003)). Geo-economic and political considerations (which Agawu emphasises) are an important factor, but the ease with which Western tonal-harmonic music has been adopted by non-Western cultures suggests that such music chimes with underlying psychoacoustic (i.e., universal) principles; that such music may be (one type of) a natural system.

3. Leman's research (2000) has shown that short-term memory plays a significant role in the perception of tonality and it "refutes the claim that probe-tone experiments provide evidence for a long-term memory of tonal hierarchies" (p. 506).

2.2.6. The Dialectical Tonic—Functionalism

Riemann proposed that tonal harmony has three functional prototypes ("harmonic pillars"), tonic (T), dominant (D), and subdominant (S). These prototypes are represented by the primary chords I, V, and IV, respectively, but they can also be represented by alterations of these chords. For example, ii is the relative minor of IV, and both are characterised as having subdominant function, thus allowing the progressions ii→V→I and IV→V→I to be similarly cadential.

Riemann's theory of tonicisation, therefore, rests upon two axioms: that cadential function is expressed by IV→V→I (S→D→T), and that a small subset of alterations (parallel, relative, and leading-tone exchange) preserve functional identity. There is good evidence to suggest that this ordering of functions is a regularity of tonal-harmonic music—it is much more common than D→S→T, particularly in cadences (Dahlhaus, 1990, pp. 57–58). There is also good evidence that different chords can share similar functionality—for example, the "subdominant" chords IV, ii, and iv; the "tonic" chords I, vi, and i.

There are two major problems with the theory. The first is epistemological; Riemann's explanation for the standard cadential order of S→D→T is that "the tonic is thetic, the subdominant antithetic, and the dominant synthetic" (as quoted in Dahlhaus (1990, p. 52)), and so most naturally follow that particular order. But there is no explanation for the correspondence of each of the musical functions with the dialectical functions (e.g., why is S, rather than D or T, antithetic?). It also assumes that the dialectical process is a fundamental part of how we cognise the world (or at least the musical part of the world), an argument that can only be made within the domain of philosophy, not science.

The second issue is theoretical; although there is a functional similarity between some common-tone transforms (like iv and IV), there is a functional dissimilarity between others (such as v and V), and so a theory of common-tone transformations preserving functionality does not provide an effective explanation for functional similarity.

Perhaps the most useful part of Riemann's theory is that it identifies and highlights important features and regularities of tonal music that require explanation: the functional similarity of certain chords within a key—e.g., the subdominant function of IV, ii, and iv, and the tonic function of I, vi, and i; the standard directionality of cadences (i.e., S→D→T, rather than D→S→T); the analytic effectiveness of a theoretical system that postulates three tonal functions; certain major-minor functional dualisms, such as the tonic-strengthening functions of the non-diatonic iv in major and its dual the non-diatonic V in minor—see Harrison (1994, pp. 15–34) for more examples. It may be this power to identify and highlight regularities of harmonic tonality that accounts for functional theory's popularity in pedagogy today, as well as continuing theoretical interest in it (e.g., Harrison (1994), Agmon (1996), Quinn (2005), and Kelley (2006)).

2.3. An Ideal Theory of Harmonic Tonality

The preceding subsections show that there is, to date, no theory that can convincingly explain the effects of tonality. But what would an ideal theory of harmonic tonality look like, and what questions should it answer?

An ideal theory of harmonic tonality should be based upon verifiable axioms that are logically developed into a theory or model that can make testable predictions of the effects it induces, explain its regularities, and shed light upon its historical development. Amongst those regularities and historical features are: (a) cadences—certain chord progressions are typically used to provide a sense of closure; (b) tonal functionality—the similar feelings of expectancy or resolution induced by some chords related by some common tone transforms; (c) tonal asymmetries—the order in which chords or keys are presented moderates their function; (d) tonal dualism—the functions of tones and chords in the major key often have a structure that is a reflection of those in the minor; (e) tonal scales—from the seven possible modes of the diatonic scale, the Ionian and Aeolian have been privileged; (f) tonality and harmony—triadic harmony and tonality developed simultaneously, suggesting they may be interdependent; (g) tonal robustness—the expectations and resolutions induced by harmonic tonality seem to be robust despite the use of different tunings and instrumental timbres.

One of the great advantages of a psychoacoustic approach is that its axioms can be experimentally tested and can use clearly defined concepts. In the next section, I will provide a full description of a novel psychoacoustic model, and I hope that it fulfils at least some of the requirements of this "ideal" theory. For the purpose of this thesis, I will focus my attention upon cadences but, in Section 4, I will briefly consider all the above regularities in more detail and see if the model can shed any light upon them as well.

3. THE PSYCHOACOUSTIC MODEL

This section presents a psychoacoustically derived computational model (built in MATLAB) of the perceived distance between any two major or minor triads, the degree of activity created by them, and the cadential effectiveness of three-triad progressions.

The underlying theory assumes the presence of six latent variables, which may be thought of as psychoacoustic or cognitive components of a listener's auditory system. The model contains a simulation of each of these latent variables and their interactions.

The psychoacoustic variables *pitch distance* (*pd*), *fundamental response distance* (*frd*), and *spectral response distance* (*srd*) are a function of psychoacoustic data (tone frequency, timbre, and frequency difference limens) and can be thought of as different metrics to determine the distance (level of difference) between any two chords (or tones). These three variables determine the value of the cognitive variables *voice-leading distance* (*vld*), and *spectral distance* (*sd*), which in turn determine the value of *tonal activity* (*act*).

The precise relationships between these six latent variables are summarised in the path diagram of Figure 3.1.

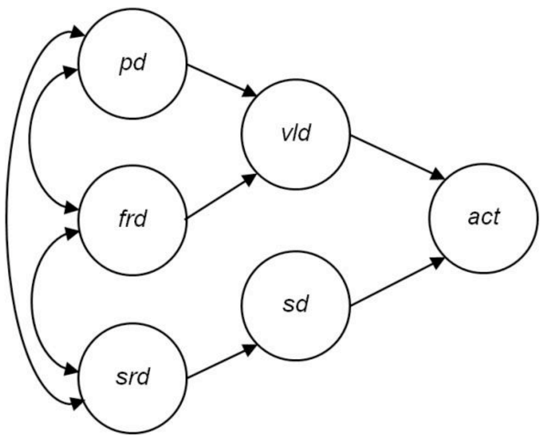


Figure 3.1. A path diagram showing the proposed flow of causation from the psychoacoustic variables pitch distance (*pd*), fundamental response distance (*frd*), and spectral response distance (*srd*), to the cognitive variables of voice-leading distance (*vld*), spectral distance (*sd*), and tonal activity (*act*). Error terms are not shown.

The model described in this thesis attempts to replicate each of these latent variables, and the relationships between them. The psychoacoustic metrics (*pd*, *frd*, and *srd*) are fully explained in Section 3.1, and the cognitive variables (*vld*, *sd*, and *act*) in Section 3.2. In Section 3.3, I discuss how the model can be used to explain and analyse harmonic cadences, while in Section 3.4, I show some ways of graphically depicting, and simplifying, the results of the model in order to facilitate its interpretation.

In the following sections, I will use the following terms and mathematical notations for chords, tones and their partials. A *partial* is a sine wave with a single frequency; a *complex tone* (or *tone*, for short) is a collection of partials, with different frequencies, that are bound together in some way (e.g., the partials move in pitch together, they are harmonic, they are produced by the same sound-source, their changes in volume are related to each other in a recognisable fashion, etc.); a *chord* is a collection of different complex tones.

When a complex tone comprises partials with frequencies that are integer multiples of a common *fundamental* frequency (as produced by most Western instruments), they are called *harmonics*; when a complex tone has harmonic partials, typically only one *pitch* is heard—its value being a function of the frequency of the fundamental's frequency; each harmonic is conventionally numbered according to the ratio of its frequency to the fundamental's (i.e., the fundamental is the first harmonic, the partial at twice its frequency is the second, etc.); in general, a tone's partials are indexed from lowest in frequency to highest, which means that when a set of partials are harmonic, their index numbers and harmonic numbers are identical.

Let x be a partial, $\boldsymbol{x}$ be a (complex) tone containing m partials, and X be a chord containing n tones. Let $x_{i,j}$ be the ith partial (or harmonic) of the jth tone ($\boldsymbol{x}_j$) of the chord X. Similarly, let $y_{k,l}$ be the kth partial (or harmonic) ($k = 1$ to q) of the lth tone ($\boldsymbol{y}_l$) ($l = 1$ to r) of the chord Y. The frequency and amplitude of $x_{i,j}$ are denoted $x_{\mathrm{f},i,j}$ and $x_{\mathrm{a},i,j}$, respectively. For this thesis, it will be assumed that all tones have harmonic partials, and that the indexing is ordered so that $x_{\mathrm{f},1,j} \leq x_{\mathrm{f},1,j+1}$ (i.e., tones are indexed from lowest in pitch to highest), and that, for any given j, $x_{\mathrm{f},i,j} < x_{\mathrm{f},i+1,j}$. For example, $x_{\mathrm{f},1}$ is the frequency of the lowest partial of the chord X, or the tone $\boldsymbol{x}$ (the context will make clear which of the two); and $y_{\mathrm{a},5,3}$ is the amplitude of the fifth harmonic of the third tone in the chord Y.

3.1. The Model's Psychoacoustic Variables

3.1.1. Pitch Distance (*pd*)

The pitch distance between two tones is approximated by the logarithm of their pitch ratio. I also assume that intervals are equivalent after octave reduction (e.g., a perfect twelfth is equivalent to a perfect fifth) and inversion (e.g., a perfect fifth is equivalent to a perfect fourth); so all intervals are no greater than six equally tempered semitones. If two tones $\boldsymbol{x}$ and $\boldsymbol{y}$ have harmonic spectra (so their virtual pitches are a function of the frequencies of their fundamentals $x_{\text{f},1}$ and $y_{\text{f},1}$), the pitch distance between them can be modelled accordingly:

$$\text{pd}(\boldsymbol{x},\boldsymbol{y}) = \min\Big(\big|\{\log_2\big(x_{\text{f},1}/y_{\text{f},1}\big)\}\big|,\ 1 - \big|\{\log_2\big(x_{\text{f},1}/y_{\text{f},1}\big)\}\big|\Big), \tag{3.1}$$

where $\{\,\}$ denotes the fractional part (i.e., $\{x\} = x - \text{int}(x)$).

When considering the pitch distance between two chords (each comprising many separate tones), I assume the overall pitch distance can be modelled by using a sum of the pitch distance moved by each voice. Each voice is also individually weighted (by the n-tuple A_j) to allow more importance to be given to, for example, the bass or soprano voices because these may be more salient. This is expressed in Equation (3.2), which assumes that both chords have the same number of voices (n), and that the voices do not cross, so the lowest voice of X (i.e., $\boldsymbol{x}_1$) is compared to the lowest voice of Y (i.e., $\boldsymbol{y}_1$), the next higher voice of X (i.e., $\boldsymbol{x}_2$) is compared to the next higher voice of Y (i.e., $\boldsymbol{y}_2$), and so on:

$$\text{pd}(X,Y) = \sum_{j=l=1}^{n} \big(A_j \times \text{pd}(\boldsymbol{x}_j,\boldsymbol{y}_l)\big). \tag{3.2}$$

Pitch distance (in conjunction with the fundamental response distance discussed later) is intended to give an indication of the voice-leading distance between two chords (see Sect. 3.2.1).

3.1.2. Fundamental and Spectral Response Distances (*frd* and *srd*)

The two response distance measures are novel metrics based upon the tenets of signal detection theory. Given a signal with a specific frequency, the auditory system is assumed to produce an *internal response* (*ir*) that may be characterised as consisting of both signal plus noise; furthermore, the noise component is assumed to have a Gaussian distribution. So the internal response to a sine wave with a specified frequency may be characterised as a Gaussian centred on that frequency. It is this noise component that makes the *frequency difference limen* (*frequency DL*) greater than zero—that is, when two sine tones of similar frequency are played successively, the listener may, incorrectly, hear them as having the same pitch.

In a two-alternative forced-choice (2-AFC) experiment, the frequency DL is normally defined as the value at which the true positive and false positive rates correspond to a d' (also known as d-prime) of approximately one. Because d' is defined as the distance between the means of two distributions divided by their standard deviation, the standard deviation of the internal response is equal to the frequency DL. This means that the internal response, as a function of frequency f, produced by a sine tone x with a frequency x_f is:

$$\mathrm{ir}_x(f) = \mathrm{e}^{-\frac{(f-x_\mathrm{f})^2}{2\mathrm{DL}(x_\mathrm{f})^2}}, \tag{3.3}$$

where $\mathrm{DL}(x_\mathrm{f})$ is the frequency difference limen at x_f. This equation allows the internal frequency response to a sine tone to be modelled using experimentally obtained measurements of frequency DLs, such as those obtained by Moore, Glasberg, and Shailer (1984).

The *response distance* (*rd*) between any two sine tones is the distance between their (Gaussian) internal responses. Although there may be many suitable metrics to measure this distance, I have chosen cosine distance because it is relatively easy to express in functional form, and because it makes intuitive sense—being the normalised cross-correlation between the two Gaussians. (A possible alternative metric would be the area under the ROC curve produced by two such Gaussian distributions.)

The cosine distance between the response curves $\mathrm{ir}_x(f)$ and $\mathrm{ir}_y(f)$, produced by two sine tones x and y with frequencies of x_f and y_f, is given by

$$\mathrm{rd}_{\cos}(x,y) = 1 - \frac{\sqrt{2}\mathrm{e}^{-\frac{(x_\mathrm{f}-y_\mathrm{f})^2}{2(\mathrm{DL}(x_\mathrm{f})^2+\mathrm{DL}(y_\mathrm{f})^2)}}}{\sqrt{\left|\frac{\mathrm{DL}(x_\mathrm{f})^2+\mathrm{DL}(y_\mathrm{f})^2}{\mathrm{DL}(x_\mathrm{f})^2\,\mathrm{DL}(y_\mathrm{f})^2}\right|\left|\mathrm{DL}(x_\mathrm{f})^2\,\mathrm{DL}(y_\mathrm{f})^2\right|}} \tag{3.4}$$

(a full derivation of this equation is provided in Appendix B).

Given two sine tones with independent frequencies, their response distance gives an indication of the probability that they are distinguishable: when they are identical in frequency, their response distance is zero; when they are far apart in frequency, their response distance approaches unity. Response distance, therefore, models a cognitive simplification of pitch distance—the former requires only a simple (binary) categorisation value of "same" or "different", while the latter requires the mind to hold a specific pitch distance value.

3.1.2.1. Spectral Response Distance

Given two successive complex tones, or chords comprising a number of complex tones, I define the *spectral response distance* to be the sum of the response distances between all possible pairings of partials where each pair contains a partial from the first chord (or tone) and a partial from the second chord (or tone). The partials in any given complex tone may have different amplitudes (typically the higher the harmonic number the lower its amplitude), so the product of their respective amplitudes weights the cosine distance for any given partial pair. For two chords X and Y, the first with m partials of frequency $x_{\mathrm{f},i}$ and amplitude $x_{\mathrm{a},i}$ (for $i = 1$ to m), the second with q partials of frequency $y_{\mathrm{f},k}$ and amplitude $y_{\mathrm{a},k}$ (for $k = 1$ to q), the total spectral response distance can be expressed accordingly:

$$\mathrm{srd}(X, Y) = \sum_{i=1}^{m} \sum_{k=1}^{q} x_{\mathrm{a},i} y_{\mathrm{a},k} \, \mathrm{rd}_{\cos}\big(x_{\mathrm{f},i}, y_{\mathrm{f},k}\big). \tag{3.5}$$

Spectral response distance is intended to give a measure of the perceived spectral distance between two chords. The values calculated by Equation (3.5) can be normalised by subtracting the value produced for two identical chords to give a minimum possible distance of zero. Also note that *srd* has a domain spanned by $m \times q$ dimensions, because every possible pair of partials between both chords is entered into it.

According to Moore, Glasberg, and Shailer (1984), the frequency DLs for harmonics within a complex tone vary according to their harmonic number (harmonics lower than five typically have a frequency DL of approximately 0.5%, harmonics higher than seven typically have a frequency DL of approximately 3%). At the time of writing, the psychoacoustic model does not allow for different widths to be chosen for different harmonics, so a compromise value of ERB/13, which corresponds to a frequency DL of approximately 1%, was chosen (ERB denotes the equivalent rectangular bandwidth, and has the value $\mathrm{ERB}(f) = 0.108f + 24.7$ (Glasberg & Moore, 1990)).

For this thesis, all the values of *srd* have been calculated with tones comprising 32 harmonic partials with amplitudes of $1/i$, where i is the harmonic number. The resulting tones, therefore, approximate sawtooth waves, and are spectrally similar to the steady-state timbre of many Western instruments (e.g., bowed strings and brass).

Furthermore, when calculating the spectral response distance between two triads, one of the triads is repeated over five octaves (two below, two above) in order to provide results that generalise better over the broad range of pitches used in music, as well as to simulate octave and inversional equivalence. For example, the progression from the note c to the note g could be down a perfect fourth (or eleventh, or eighteenth) or up a perfect fifth (or twelfth or nineteenth), and so forth.

3.1.2.2. Fundamental Response Distance

The response distance can also be applied to just the fundamentals of each tone. This *fundamental response distance* (in conjunction with pitch distance) is intended to give an indication of the voice-leading distance between two chords. Let X be a chord containing n tones with harmonic partials; the frequency and amplitude of its tones' fundamental partials are denoted $x_{\mathrm{f},1,j}$ and $x_{\mathrm{a},1,j}$, respectively, where j indexes each tone in the chord (from lowest to highest, and $j = 1$ to n). Similarly, let Y be a chord containing r tones with harmonic partials; the frequency and amplitude of its tones' fundamental partials are denoted $y_{\mathrm{f},1,l}$ and $y_{\mathrm{a},1,l}$, respectively, where l indexes each tone in the chord (from lowest to highest, and $l = 1$ to r). This means that the fundamental frequencies and amplitudes of each tone in X are $x_{\mathrm{f},1,1}, x_{\mathrm{f},1,2}, \ldots, x_{\mathrm{f},1,n}$ and $x_{\mathrm{a},1,1}, x_{\mathrm{a},1,2}, \ldots, x_{\mathrm{a},1,n}$; the fundamental frequencies and amplitudes of each tone in Y are $y_{\mathrm{f},1,1}, y_{\mathrm{f},1,2}, \ldots, y_{\mathrm{f},1,r}$ and $y_{\mathrm{a},1,1}, y_{\mathrm{a},1,2}, \ldots, y_{\mathrm{a},1,r}$. I define the fundamental response distance between the chords X and Y to be

$$\mathrm{frd}(X,Y) = \sum_{j=1}^{n}\sum_{l=1}^{r} x_{\mathrm{a},1,j}\, y_{\mathrm{a},1,l}\, \mathrm{rd}_{\cos}\big(x_{\mathrm{f},1,j}, y_{\mathrm{f},1,l}\big). \tag{3.6}$$

Note that the domain of *frd* is spanned by $n \times r$ dimensions, because every possible pair of fundamentals between both chords is entered into the calculation.

According to Moore, Glasberg, and Shailer (1984), the frequency DL for a complex harmonic tone, as a whole, is smaller than that for any of its partials, and generally approximates 0.2%. This is approximated by ERB/66, which is the value used in the model to calculate fundamental response distance.

The Gaussian noise component of the internal frequency response is, therefore, relatively narrow compared to the smallest musical interval used in common practice (the semitone). This means that the fundamental response distance effectively acts as a counter for the number of non-common tones between two chords. That is, it gives a distance of zero to two identical triads, a distance of approximately one third to two triads with two common tones (e.g., parallel triads like C and c, relative triads like C and a, leading tone exchange triads like C and e), a distance of approximately two thirds to two triads sharing one common tone (e.g., dominant triads like C and G), and a distance of approximately unity to two triads with no common tone (like C and D). This is clearly in accord with Riemannian and neo-Riemannian music theory, which treats the above-mentioned common-tone transformations as being especially close (see, e.g., Kopp (2002)).

3.2. The Model's Cognitive Variables

The psychoacoustic variables of the model are assumed to directly affect the cognitive variables. However, the paths from psychoacoustic to cognitive are mediated by musical (and long-term memory) constraints; given a (familiar) musical system with a specific structure or set of rules—such as the use of instruments with approximately harmonic timbres, and a limited number of differently tuned tones—only certain spectral and voice-leading possibilities are actually possible (expected).

Let it be assumed that the musical system under analysis uses a number of independent tones each of which consists of dependent (i.e., approximately harmonically related) spectra. This is a fair description of the majority of Western music, where voices move with some degree of independence, and the majority of these voices are harmonic complexes with a clear sense of pitch. This type of musical system creates strict constraints upon movements within the continuum of all possible spectral tunings. For example, imagine we are able to create any possible spectrum, containing 16 independently tuned partials, and let any specific point in this 16-dimensional spectral continuum be denoted a *spectral tuning*; it would be possible to move from any arbitrary spectral tuning within this space to any other. But in conventional music, with the above-mentioned constraints, we can control—and are accustomed to hearing—the movement of a limited number of voices (typically four) built from tones comprising harmonic partials. This means that the range of musically possible spectral tunings, and possible paths between them, is substantially constrained.

Voice-leading distance (*vld*) is the cognitive distance between two spectral tunings under these musical constraints. The distances along these constrained paths are mediated by the pitches (fundamentals) of each tone; so it makes sense to hypothesise that voice-leading distance is a function of pitch distance and fundamental response distance (because, these two distances are concerned only with the frequencies of the fundamentals). Voice-leading distance can be used to describe the musically-constrained distance between any two triads; it can also be used to describe the musically-constrained distance between two pairs of triads: given two pairs of triads, $U \leftrightarrow V$, and $X \leftrightarrow Y$, the voice-leading distance between them is equivalent to the voice-leading distances between U and X plus the voice-leading distance between V and Y (or the voice-leading distance between U and Y plus the voice-leading distance between V and X, whichever is smaller).

Spectral distance (*sd*), on the other hand, is the unconstrained cognitive distance between any two spectral tunings. Because it is a function of the tuning of all partials in each chord, it makes sense to hypothesise it to be a function of the spectral response distance.

A corollary of having two independent distances is that it is possible for a pair of triads to be voice-leading close but spectrally distant; or for a pair of triads to be voice-leading distant but spectrally close. This has a very important consequence: given two pairs of chords that are voice-leading close (e.g., the two triad pairs D♭↔G and D♭↔g are voice-leading close because G and g differ by just one semitone in one voice), such that one pair is spectrally more distant than the other (in reference to the triad D♭, the triad G is spectrally more distant than the triad g), the spectrally distant pair will tend to be heard as if it were a voice-leading alteration of the spectrally closer pair. I hypothesise that this sense of alteration (i.e., of a more “complex”, or “difficult”, choice made as a substitute for a “simpler”, more “straightforward”, choice) is the origin of *tonal activity*, or expectation. I also hypothesise that this activity is resolved by allowing the altered tone to move to a different tone that is in a triad that is spectrally close to (has a “simple”, “straightforward” relationship with) preferably both of the preceding two chords.

The next three subsections discuss voice-leading distance, spectral distance, and tonal activity in more detail.

3.2.1. Voice-Leading Distance (*vld*)

Voice-leading distance is the latent variable representing the cognitive distance between the (musically constrained) voices of two chords, or two pairs of chords. When assessing the perceived distance between two chords, it is common to measure the overall pitch distance between them. This is typically calculated as the city block, Euclidean or other Minkowski, distance between the semitone values of the notes in two chords. It seems reasonable to assume this is a good measure for pairs of tones, or other simple stimuli. But when it comes to measuring the distance between triads, or between any voice-leading involving three or more parts, is it reasonable to expect a listener to individually track the degree of movement of every voice before summing them?

As discussed above, for standard musical tuning systems, the fundamental response distance is effectively binary—it has a value of zero for a common tone, a value of one for anything else, and so counts the number of non-unisons between any two triads. It seems plausible that, due to the simplicity of this (neo-Riemannian-like) binary measure, the fundamental response distance may also play a part in determining the voice-leading distance for more complex stimuli (such as three, or more, part voice-leadings).

We might expect a listener, therefore, to judge the voice-leading distance between two chords to be a function of pitch distance and fundamental response distance. Furthermore, we might expect the pitch distance between the more salient voices, such as the lowest, highest, or root, to be more important than the pitch distances between less salient voices.

For two chords X and Y both containing n tones indexed by j and l (so X contains the tones $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$; Y contains the tones $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n$), voice-leading distance is modelled in the following way:

$$\text{vld}(X, Y) = \sum_{j=l=1}^{n} \big(A_j \times \text{pd}(\boldsymbol{x}_j, \boldsymbol{y}_l)\big) + \big(B \times \text{frd}(X, Y)\big). \tag{3.7}$$

Observe that the first part of the right-hand side is the city block pitch distance between tones of the same voice, except that each voice is individually weighted. This individual weighting allows more importance to be given to, for example, the bass or soprano voices because these may be more salient (this part of the equation assumes that voice crossing does not occur). The domain of this part of the formula is spanned by n dimensions, because only pairs with matching voice indexes are entered. The second part of the right-hand side "counts" the number of non-common tones regardless of their voice index numbers. The domain of this part of the formula is spanned by n^2 dimensions because all possible pairs of the tones' fundamentals (where one tone is in chord X and one is in chord Y) are used (see Sect. 3.1.2.2). In Section 5.2, empirical data is used to determine values for the parameters A_j and B (see Eq. (5.1)).

The voice-leading distance between two pairs of chords ($U \leftrightarrow V$ and $X \leftrightarrow Y$), all containing n voices, is calculated accordingly:

$$\text{vld}(U \leftrightarrow V, X \leftrightarrow Y) = \min\big(\text{vld}(U, X) + \text{vld}(V, Y),\ \ \text{vld}(U, Y) + \text{vld}(V, X)\big). \tag{3.8}$$

Note that when U OR $V = X$ OR Y, the domain is still spanned by n^2 dimensions.

3.2.2. Spectral Distance (*sd*)

Spectral distance is the latent variable representing the cognitive distance between the spectra of two chords. In the same way it seems unreasonable to expect a listener to track the movement of every single tone in a three, or more, part voice-leading, it is even more unreasonable to expect a listener to track the distance moved by every single partial found in one chord to the partials found in a second chord. Furthermore, in normal listening even those partials that can be resolved are not actually "heard out"; instead they are subsumed into the unified perceptions of virtual pitch and timbre. Furthermore, even if they were actively heard out, it would be almost impossible to know in which direction any given partial "moves"—does it "go" to the partial that is closest in pitch, or the partial that has the same position in a frequency-ranked stack of partials? The spectral response distance does not attempt to "track" any supposed motion of partials, it simply scores every coincident pair as zero, every non-coincident pair as unity. Every pair that is almost coincident is given a score between zero and unity—its precise value determined by the width of the underlying Gaussian internal response curve.

For this reason, it is reasonable to expect the spectral distance (*sd*) between any two triads to be strongly correlated to their spectral response distance (*srd*) and, in the current model, they are treated as mathematically equivalent (i.e., $\text{sd}(X, Y) = \text{srd}(X, Y)$). At first sight, it may seem odd to keep *sd* and *srd* conceptually separate, but further testing may show the cognitive component (*sd*) is not linearly related to the psychoacoustic component (*srd*), and that there may be other, as yet unidentified, components that also determine *sd*. In other words, a more developed model may define spectral distance accordingly: $\text{sd}(X, Y) = f(\text{srd}(X, Y)) + \zeta$, where f is a monotonic function, and ζ represents currently unidentified components of the model (e.g., long-term memory and neurological priming may impact upon our cognition of spectral distance). For that reason, *srd* and *sd* are kept conceptually separate, even though there is, at this stage, no mathematical requirement.

In tonal terms, spectral distance acts as a type of weighted counter because a smaller distance is given to those intervals (dyads) whose tunings approximate low integer frequency ratios (such as the perfect fifth, which approximates 3/2). This is because low-integer (simple) ratio intervals of tones with harmonic spectra have more coincident partials. For harmonic tones, the ratio of coincident to non-coincident partials can be expressed mathematically, which enables the *approximate spectral response distance* between two tones, $\boldsymbol{x}$ and $\boldsymbol{y}$, with harmonic partials to be easily calculated by hand:

$$\sim\text{srd}(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{1}{2s} - \frac{1}{2t}, \tag{3.9}$$

where s/t is the frequency ratio, in reduced form (i.e., s and t are coprime), of $\boldsymbol{x}$ and $\boldsymbol{y}$ (see Appendix B for the derivation of this equation). Equation (3.9) shows that melodic dyads conventionally considered harmonically consonant (perfect fifths and fourths, and thirds and sixths, which typically approximate simple ratios) have lower spectral distance than dyads like seconds and sevenths and the tritone, which approximate more complex ratios.

When considering the total spectral distance between two triads, the spectral distance of every single melodic dyad between them (there are nine melodic dyads between the tones of two triads) can be separately calculated, by hand, using Equation (3.9), and then summed. This means that the greater the number of dyads between two triads that approximate simple ratios (i.e., they have low-valued s and t), and the greater their simplicity, the lower the spectral response distance between them.

This approximation has been introduced only to provide a simple illustration of the relationship between the complexity of a dyad's frequency ratio and its spectral distance. The psychoacoustic model uses the full calculation, not this approximation.

3.2.3. Tonal Activity (*act*)

I hypothesise that tonal activity is the result of the interplay between voice-leading distance (which is a function of pitch distance and fundamental response distance) and spectral distance (which is a function of spectral response distance).

Given a musically presented triad pair, let any other pair against which that musically presented pair is mentally compared, be denoted a *comparison* pair. I hypothesise that when a triad pair has a higher spectral distance than a comparison triad pair that is voice-leading close, the former triad pair may be heard as an alteration of that comparison triad pair. This sense of alteration creates a feeling of "activity" and, for two comparison triad pairs, the activity is greater when the difference between their spectral response distances is greater, and the voice-leading distance between them is smaller.

This can be stated more formally: Let there be a pair of played triads, $X \leftrightarrow Y$, that are mentally compared to another pair, $U \leftrightarrow V$, that is held in memory. The tonal activity of $X \leftrightarrow Y$, given the (cognitive) existence of the comparison pair $U \leftrightarrow V$, is denoted $\mathrm{act}(X, Y \,|\, U, V)$, and is calculated accordingly:

$$\mathrm{act}(X, Y \,|\, U, V) = f\left(\frac{\mathrm{sd}(X, Y) - \mathrm{sd}(U, V)}{\mathrm{vld}(U, X) + \mathrm{vld}(V, Y)}\right), \qquad (3.10)$$

where f is a monotonic function (see the next two subsections for a description of two possible monotonic functions). In the subsequent analyses, I will be concerned mostly with comparison pairs where only one of the triads is different: that is, the tonal activity of $U \leftrightarrow Y$, or $X \leftrightarrow V$, due to the comparison pair $U \leftrightarrow V$. In such cases, the above equation simplifies to $\text{act}(U, Y \mid U, V) = f\left(\frac{\text{sd}\,(U,Y) - \text{sd}\,(U,V)}{\text{vld}\,(V,Y)}\right)$, or $\text{act}(X, V \mid U, V) = f\left(\frac{\text{sd}\,(X,V) - \text{sd}\,(U,V)}{\text{vld}\,(U,X)}\right)$, respectively.

In the next two subsections (3.2.3.1 and 3.1.2.2), I propose two monotonic functions, each giving separate (but correlated) measures of tonal activity: *continuous activity* (act_c), and *discrete activity* (act_d). Both these measures are experimentally tested in Section 5.2, and I assume that *tonal activity* (act) is a linear combination of both continuous and discrete activity though, due to their high collinearity, it is not currently possible to determine their relative importance and only the "best" out of these two measures is used in the regression equations of Section 5.2.

3.2.3.1. Continuous Activity (act_c)

The continuous activity measure simply assumes the monotonic function in Equation (3.10) to be linear:

$$\text{act}_c(X, Y \mid U, V) = \frac{\text{sd}(X, Y) - \text{sd}(U, V)}{\text{vld}(U, X) + \text{vld}(V, Y)}. \tag{3.11}$$

A result of this definition is that $\text{act}_c(X, Y \mid U, V) = -\,\text{act}_c(U, V \mid X, Y)$. Assuming the absolute level of continuous activity is more than negligible, I hypothesise that the triad pair with positive continuous activity is heard as an alteration of the comparison pair with negative continuous activity. This is because the two triads in a pair with high positive continuous activity have a more distant (complex) spectral relationship than the voice-leading close pair with negative continuous activity. I presume that triad pairs with negative continuous activity tend to sound "passive", "stable", and "at rest", while triad pairs with positive continuous activity tend to sound "active", "unstable", and "restless".

Let me illustrate this concept with a relatively straightforward example. The psychoacoustic model predicts that the two root-position triad pairs C↔d and C↔D are voice-leading close (D and d have two common tones, one of which is the salient bass note). It also predicts the two triads C↔d are spectrally closer than the two triads C↔D (the latter pair replaces the former's low spectral distance perfect fourth—between the root of the first chord and the third of the second chord—with a high spectral distance tritone). So the former pair has negative continuous activity, the latter has positive continuous activity, which means that the latter is heard as an alteration of the former.

(If that example still seems difficult to understand, consider just two successive tones. We might consider melodic intervals of a semitone and tritone to be active because they can be mentally compared to the voice-leading close melodic intervals of the unison and perfect fourth/fifth, respectively. The process for triads, described above, is simply an extension of this concept to a higher-dimensional tone space—illustrations of which are given in Sect. 3.4.)

3.2.3.2. Discrete Activity (act_d)

The discrete activity measure represents a cognitive simplification of the continuous measure. It models the probability that a triad pair will be discretely categorised as being either active (an alteration of a comparison pair) or inactive (not an alteration of a comparison pair).

This process of representing a continuous variable with a simpler categorical variable is very similar to how the response distance of two sine tones represents a cognitive simplification of their pitch distance (see Sect. 3.1.2): pitch distance is a continuous measure of the distance between two sine tones; response distance is the probability the two tones will be categorised as either "different" or "the same". I presume that a similar process of cognitive simplification can also occur for activity and, under cognitive load (and maybe as mediated by long-term memory), activity is cognitively categorised as being either present or not: its precise magnitude being cognitively ignored or discarded.

Discrete activity is defined as the probability of a triad pair being heard as active, due to a given comparison, and it can be simply modelled with a logistic function of that triad pair's continuous activity:

$$\mathrm{act_d}(X, Y \mid U, V) = \frac{1}{1 + e^{-k \times \mathrm{act_c}(X, Y \mid U, V)}}; \tag{3.12}$$

so as a triad pair's continuous activity decreases below zero, its discrete activity tends towards zero; as its continuous activity increases above zero, its discrete activity tends towards unity. Attempting to empirically derive the value of the parameter k (which controls the width of the transition) is beyond the scope of this thesis so, for simplicity's sake, I assume it approaches infinity. This means the above function can be replaced by the Heaviside step function, $H_0(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$, that is,

$$\mathrm{act_d}(X, Y | U, V) = H_0(\mathrm{act_c}(X, Y | U, V)). \tag{3.13}$$

3.2.3.3. Asymmetry of Tonal Activity

When considering any given type of comparison, the model of tonal activity has a plausible asymmetry. For the following explanation and examples, I will consider just the parallel comparison, e.g., C↔D compared with C↔d; or c↔D compared with C↔D; or c↔D compared with C↔d. The reason for privileging the parallel comparison is because it is reasonable to surmise that, for root-position triads, the parallel transformation will be judged to have the smallest voice-leading size (Sect. 5.2.3 also provides experimental support for this). This is because it has two common tones; the non-common tone moves the smallest possible pitch distance (one semitone); the non-common tone is not the salient root (bass note) of the two chords. The parallel transform is the only one that has all three of these characteristics. Equation (3.10) implies that the absolute value of activity is maximised by having a comparison pair that is voice-leading close (i.e., both $\mathrm{vld}(U, X)$ and $\mathrm{vld}(V, Y)$ are small), so by choosing the parallel comparison, we are likely to be exploring those tonal activities that are most important to our perception of music (see Sect. 5.1.3.2 for a further discussion of this issue).

For more compact notation, let the activity of the pair $X \leftrightarrow Y$, where X is played first and Y second, due to comparison with its parallel transform pair $X \leftrightarrow V$, be denoted $\mathrm{act}(X \rightarrow Y | \mathbf{P})$; the arrow points rightwards to the second triad because this triad is being compared to its transform, and the bold letter indicates the type of transform: **P** is the parallel transform, though **R**, **L**, **S**, and **S** + **P**, etc. could be used to denote the relative, leading tone exchange, SLIDE (e.g., C↔c♯) (Lewin, 1987), and SLIDE + parallel (e.g., C↔C♯) transforms, respectively. Similarly, the activity of the pair $X \leftrightarrow Y$, where X is played first and Y second, due to comparison with its parallel transform pair $U \leftrightarrow Y$ is denoted $\mathrm{act}(X \leftarrow Y | \mathbf{P})$ (in this case, the arrow points leftwards to the first chord, because this chord is being compared to its parallel transform). Activities with a rightwards arrow are called *forwards activities*; activities with a leftwards arrow are called *backwards activities*. In the absence of any broader context, it is reasonable to assume that forwards activities are more salient than backwards activities, because the most recently played (or currently playing) triad will have greater prominence in the listener's short-term or echoic memory.

Generally, $\mathrm{act}(X \rightarrow Y|\mathbf{P}) \neq \mathrm{act}(X \leftarrow Y|\mathbf{P})$. For example, the model calculates that $\mathrm{act}_\mathrm{c}(\mathrm{C} \rightarrow \mathrm{D}|\mathbf{P}) < 0$, while $\mathrm{act}_\mathrm{c}(\mathrm{C} \leftarrow \mathrm{D}|\mathbf{P}) < 0$ (which implies that $\mathrm{act}_\mathrm{d}(\mathrm{C} \rightarrow \mathrm{D}|\mathbf{P}) = 1$, while $\mathrm{act}_\mathrm{d}(\mathrm{C} \leftarrow \mathrm{D}|\mathbf{P}) = 0$). In words, given the pairing of triads C and D, the D is heard as altered, rather than the C. A natural consequence of is that a putative cadential progression proceeding from C to D to some resolution triad is likely to be more effective than a progression from D to C to some resolution triad (because, in the former case, the active triad is more prominent in the listener's short term or echoic memory). Hence the temporal asymmetries of tonality (e.g., see Dahlhaus' discussion of the order of "functions" within cadences (1990)) are a natural consequence of the proposed activity function. Such asymmetries cannot be explained by inherently symmetrical structural models, such as Lerdahl's (2001), without the addition of a separate layer of theory.

3.2.3.4. Exclusive and Non-Exclusive (Double) Alterations

When considering a triad pair such as C→f, it is reasonable to say that this could be heard as an alteration of either C→F or c→f, because the triads in either pair have a lower spectral distance than C→f. It is also reasonable to assume that only one of these possible alterations is likely to be heard at any given time—that is, either C is heard as altered, or f is heard as altered, but not both at the same time. The reason these two alterations are exclusive is because if both were assumed to coexist, it would imply a comparison triad pairing of c→F, which is itself an alteration ($\mathrm{act}_\mathrm{c}(\mathrm{c} \rightarrow \mathrm{F}|\mathbf{P}) > 0$ and $\mathrm{act}_\mathrm{c}(\mathrm{c} \leftarrow \mathrm{F}|\mathbf{P}) > 0$) and so cannot function as a stable reference point.

However, there are three triad pairings (c↔D, C↔e♭, and c↔E) in which both triads can be heard as simultaneously altered. Let $X \leftrightarrow Y$ be a triad pairing, and let $U \leftrightarrow V$ be a comparison pairing such that $U \leftrightarrow X$ are parallels, and $V \leftrightarrow Y$ are also parallels. Both triads X and Y can be heard as simultaneously altered if, and only if, $\mathrm{act}_\mathrm{c}(X \rightarrow Y|\mathbf{P}) > 0$ AND $\mathrm{act}_\mathrm{c}(X \leftarrow Y|\mathbf{P}) >$

0 AND $\mathrm{act}_c(U \rightarrow V|\mathbf{P}) < 0$ AND $\mathrm{act}_c(U \leftarrow V|\mathbf{P}) < 0$. Any such pairing is denoted a *double alteration*.

For example, c→D is heard as an alteration of c→d, which is heard as an alteration of C→d, so c→D is heard as a double alteration of C→d and has two active notes—*e*♭ and *f*♯ (alternatively, it could be said that c→D is heard as an alteration of C→D, which is heard as an alteration of C→d, etc., but this amounts to the same conclusion). Similarly, C→e♭ can be interpreted as an alteration of C→E♭, which can be interpreted as an alteration of c→E♭, so C→e♭ is heard as a double alteration of c→E♭ and has two active notes—*e* and *g*♭.

3.3. Applying the Model to Harmonic Cadences

I hypothesise that cadential effectiveness is a function of the tonal activities between each pair of triads in a putative cadence, and that there are four other factors that may also impact upon cadential effectiveness. All five factors are described in the following subsections.

3.3.1. Cadential Form—Patterns of Tonal Activity

When two triads are played successively and one of those triads is heard as an active alteration of the other, we expect it to resolve. For a third triad to act as a successful resolution, the pairings it makes with both the earlier triads should have low (negative) activity. If this is not the case, the third triad is creating new active notes that may require further resolution, and the cadence may feel incomplete.

This structure gives a template for harmonic cadences formed with three triads denoted *antepenult* (A), *penult* (P), and *final* (F) (i.e., the putative cadential progression is $A \rightarrow P \rightarrow F$, so there are six different activity values to be considered: $\mathrm{act}(A \rightarrow P|\mathbf{P})$, $\mathrm{act}(A \leftarrow P|\mathbf{P})$, $\mathrm{act}(P \rightarrow F|\mathbf{P})$, $\mathrm{act}(P \leftarrow F|\mathbf{P})$, $\mathrm{act}(A \rightarrow F|\mathbf{P})$, $\mathrm{act}(A \leftarrow F|\mathbf{P})$.

(The terms "antepenult", "penult", and "final" come from linguistics, where they indicate the position of syllables in a word; I have adopted these terms to describe three-triad progressions because their meaning is easy to deduce, and they are free of the functional and scale degree associations of terms like "subdominant", "dominant", and "tonic".)

I hypothesise that an effective cadence requires the pairing between A and P to have high (positive) activity, and the pairings between P and F, and between A and F, to have low (negative) activity; the classic IV→V→I and ii→V→I cadences (e.g., C→D→G and d→G→C, respectively) have precisely this pattern of activities. Generally speaking, we might expect to see a positive correlation between act$(A \rightarrow P|\mathbf{P})$ and act$(A \leftarrow P|\mathbf{P})$ and cadential effectiveness, and a negative correlation between act$(P \rightarrow F|\mathbf{P})$, act$(P \leftarrow F|\mathbf{P})$, act$(A \rightarrow F|\mathbf{P})$, and act$(A \leftarrow F|\mathbf{P})$ and cadential effectiveness.

For those triad pairs where the backwards and forwards activities are exclusive (e.g., C↔f), only the forwards activity is used (because it has greater salience—see Sect. 3.2.3.3), and the backwards activity value is replaced with zero. For those doubly active triad pairings where the forwards and backwards activities can coexist (e.g., C↔e♭), both activity values are used.

3.3.2. Cadential Fit—Spectral Distances

The spectral distances between each pair may also have a direct impact on cadential effectiveness. If the triads are spectrally distant they may "fit badly", making the progression sound "difficult", "clumsy", or "unnatural"—high spectral distance cadences may be aesthetically unappealing, or the extra cognitive load they may require may distract from their cadential function. Because the $A \leftrightarrow P$ pair should be active, the spectral distance may be less important for that particular pair.

3.3.3. Cadential Salience and Flow—Voice-Leadings Toward Resolution

The manner in which active notes resolve from one triad to the next (and most particularly from the penult to final) may impact upon the overall effectiveness of the resolution. There are three possible variations in the manner of resolution, discussed below.

3.3.3.1. Position in Final

It is generally accepted that the most salient tone of a triad is its root (Parncutt, 1988); therefore, for a resolution to be as salient as possible, active tones and resolution tones should (if possible) be roots. In this thesis, I am principally examining activities due to parallel comparisons, which means the active tone is always the third of the triad. The resolution tone, however, is not constrained—it can take any position of the triad. It is reasonable to expect, therefore, that the most effective resolutions will be those where the penult's active tone resolves to the final's root.

3.3.3.2. Direction

The direction of the resolution is likely to be an important factor in cadential effectiveness. An active tone is perceived as an alteration of another tone; it has, therefore, *directionality*. For example, in the progression C→D, the active *f*♯ is heard as an upwards, not a downwards, alteration of *f*♮. The motion towards resolution is more "directed", or aesthetically consistent, if it continues in the same direction.

3.3.3.3. Amount

The size of the voice-leading between the penult's active tone and its resolution tone in the final may also impact upon cadential effectiveness. It may be that smaller movements (e.g., semitones) induce a more effective sense of resolution than larger movements (e.g., whole tones).

3.3.4. Cadential Synergy—Latent Activities of the Embedding Scale

Chord pairs, triples, and so on, do not exist as independent units isolated from a broader context of possibilities: they exist within a *scale* (a set of notes that is bound together in some way) that is either explicitly spelled out (in full) by the chords used, or implicitly indicated. For example, a perceptually simple scale—for example, a well-formed scale like the diatonic (well-formed scales are those consisting of just two step sizes distributed as evenly as possible (Carey & Clampitt, 1989))—or a scale familiar to the listener, may be partially spelled out by the chords used, the rest of the notes being mentally "filled in" by the listener.

Most scales embed a number of possible harmonic progressions that are not necessarily actualised by a given chord progression. For example, the progression C→D contains the notes *c*, *d*, *e*, *f*♯, *g*, *a*, and this scale contains not just C and D, but also the triad a. Furthermore, it is likely that such a progression will be heard as part of the familiar well-formed diatonic scale *c*, *d*, *e*, *f*♯, *g*, *a*, *b*, which contains the additional triads e, G, and b. (From an analytical point of view, it is perhaps safest to consider only those scales that are explicitly indicated by a progression.)

This means that, when considering the overall impression of a triad progression, it is necessary to consider the scalic context, or contexts, within which the progression is embedded—that is, it is necessary to consider the latent activities of the embedding scale. For example, it may be that a putative cadence spells out a scale all of whose latent activities support the same tonic triad—there is, therefore, *cadential synergy*; alternatively, the latent activities may give contradictory tonics, so that a triad that is a tonic for one latent progression may be active in a different latent progression.

3.3.5. Cadential Complexity—Melodic Complexity of Embedding Scale

The embedding scale may be melodically simple in form (e.g., it may be well-formed, or a scale that is similar to well-formed). Alternatively, it may be melodically complex. One clear marker of a type of scale that is so complex it is infrequently used (for tonal-harmonic music) is the presence of two consecutive semitones. Tymoczko's "no consecutive semitones" rule, states that this is a necessary condition for scales to have a clear distinction between steps and leaps; for that reason, most tonal-harmonic music, even contemporary, avoids such scales (2004).

Putative cadential progressions that spell out such scales may, therefore, be heard as having excessively complex melodic latencies, and may be heard as less effective as a result.

3.3.6. The Model's Calculations

The current model produces direct calculations for the spectral response distance and tonal activity for all pairs of triads and, in Section 5, these values are used as variables in regressions of ratings of "similarity", "fit", and "cadential effectiveness". The model, therefore, calculates cadential form and cadential fit, but it does not (currently) produce calculations for cadential salience, flow, synergy, or complexity, so these form no part of the regressions analyses. However, these factors are discussed, in broad terms, in Section 4.1 (which considers the ability of the model to predict the conventional cadences used in tonal harmonic music).

3.4. Interpreting the Model

3.4.1. Plotting the Model's Calculations

The implications of the model are easier to understand and interpret when its calculations are plotted. The continuous activity (act_c) of one triad pair, as a result of comparison with another triad pair, is equivalent to the difference between their spectral distances (*srd*) divided by the voice-leading distance (*vld*) between them (see Eq. (3.5)). This means that if two triad pairs are plotted such that the voice-leading distance between them is plotted on one axis, and their respective spectral response distances are plotted on another axis, the continuous activity of each triad pair, due to comparison with the other, is equivalent to the slope between them. It is possible to produce such a plot for any two pairs of triads but, because the voice-leading metric *vld* is not Euclidean, it is impossible to do this for a continuum of different pairs of triads. As discussed in Section 3.2.1, the voice-leading metric (see Eq. (3.7)) is a function whose domain is spanned by n^2 dimensions (where both chords have n tones each), so it could be (crudely) approximated by an n^2-dimensional Euclidean metric. But that still does not enable a visually informative plot to be produced, because a minimum of nine dimensions would be required for pairs of triads. However, if the fundamental response distance (i.e., the count of non-common tones) is removed from the *vld* measure, the domain of the *vld* function is now spanned by only n dimensions; furthermore, if only root-position triads are considered, only $n-1$ dimensions are required. This is because all root-position major and minor triads have a perfect fifth above their bass, so the bass voice and the fifth voice do not move independently, which means these two dimensions can be concatenated into a single root + fifth dimension.

This means that the voice-leading distance between root-position triads, connected by parallel motion, can be (very) crudely approximated by the following two-dimensional Euclidean metric,

$$vld \sim \sqrt{2\left|\log_2(x_{f,1,1}) - \log_2(y_{f,1,1})\right| + \left|\log_2(x_{f,1,3}) - \log_2(y_{f,1,3})\right|}, \quad (3.14)$$

where $x_{f,1,1}$ and $y_{f,1,1}$ are the frequencies of the fundamentals of the two triads' roots (or fifths), and $x_{f,1,3}$ and $y_{f,1,3}$ are the frequencies of the fundamentals of the two triads' thirds.

This results in the forms illustrated in Figure 3.2—(a), (b), and (c) have a major reference triad, while (d), (e), and (f) have a minor reference triad; both reference triads are tuned to twelve tone equal temperament (12-TET). In (a) and (d), the horizontal axes show the pitch distance between the roots (and fifths) of the reference triad and a continuum of differently tuned triads, the vertical axes show the pitch distance between the thirds of the reference triad and a continuum of differently tuned triads. They are scaled as indicated by Equation (3.14), so a one semitone distance on the horizontal axis is $\sqrt{2}$ longer than a one semitone distance on the vertical axis. This ensures that all straight-line distances between triads are Euclidean. The spectral response distance between the reference triad and every continuum triad (i.e., at each point on the plot) is illustrated with a greyscale that is black at the global minimum spectral distance, gets lighter as spectral distance increases, and is white at the global maximum spectral distance.

To provide a more precise visualisation of the topology, the remaining figures view (a) and (d) from the side, so the vertical axes show the spectral response distance. In (b) and (e), the horizontal axes show the pitch distance between the roots (and fifths) of the reference triad and a continuum of differently tuned triads; in (c) and (f), the horizontal axes show the pitch distance between the thirds of the reference triad and a continuum of differently tuned triads.

I have labelled the location of a selection of major and minor triads to help orient the reader (note that in (a) and (d), the major and minor continuum triads run up the two diagonal lines, with major triads located vertically above their minor parallel). The precise values of *srd* calculated by the model for every 12-TET triad pair can be found in Appendix C.

When using these figures it should be remembered that, due to the excision of the *frd* component of *vld*, the voice-leading distances they show are approximations; it is likely that triads with common tones (e.g., C and e, or C and a) have substantially closer voice-leading distances than shown here (but, as explained above, it is not possible to graphically represent all such relationships in a low-dimensional Euclidean space).

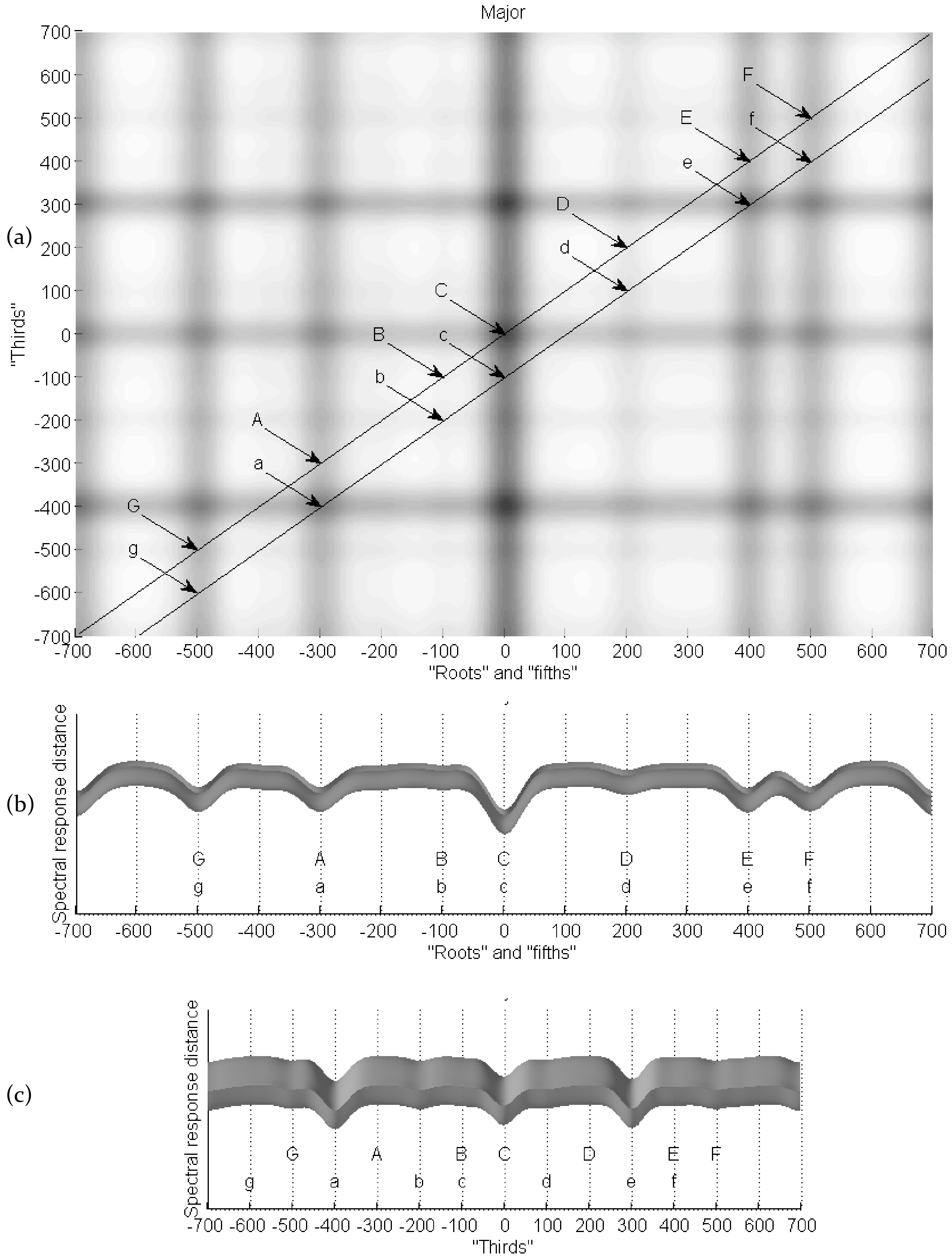
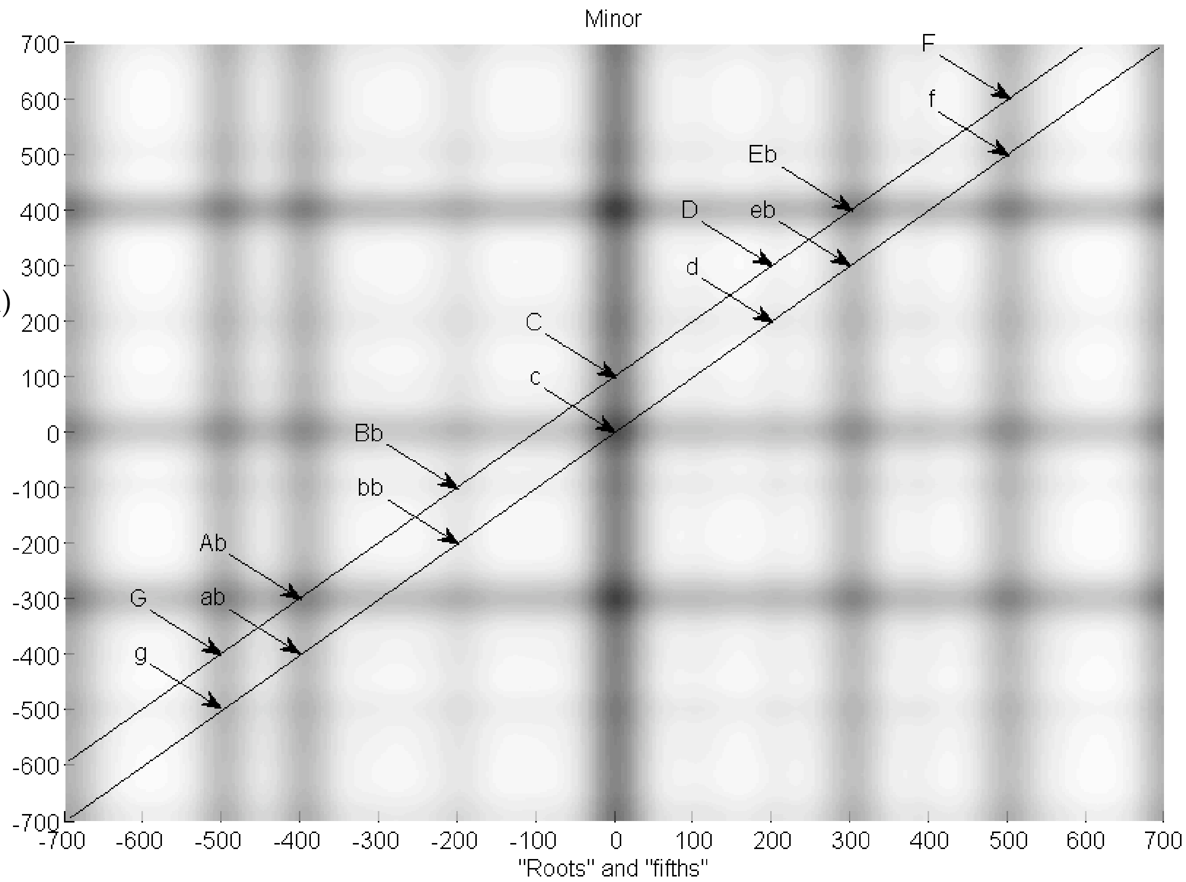
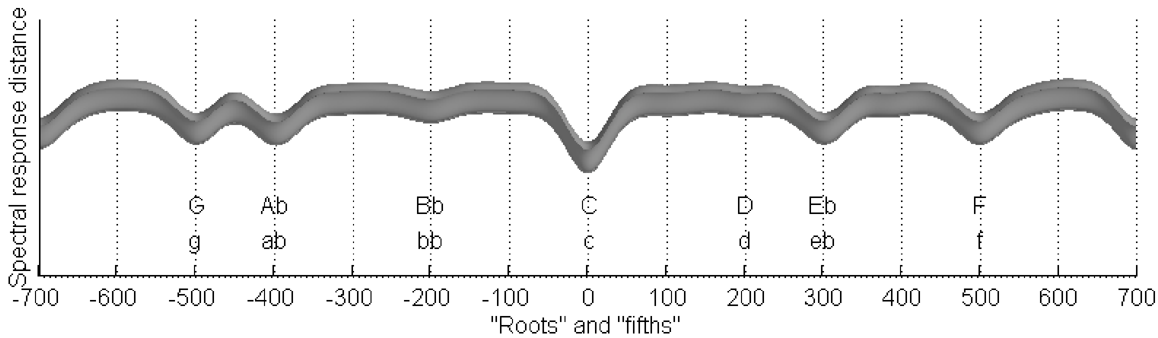


Figure 3.2. The spectral response distances between a 12-TET reference triad and a continuum of differently tuned triads, plotted against their pitch distances. The reference triad is major in (a), (b), and (c), minor in (d), (e), and (f). In (a) and (d), the horizontal axes show the pitch distance (in cents) between the "roots" and "fifths" of the continuum triads and the root and fifth of the reference triad; the vertical axis shows the pitch distance between the "thirds" of the con-

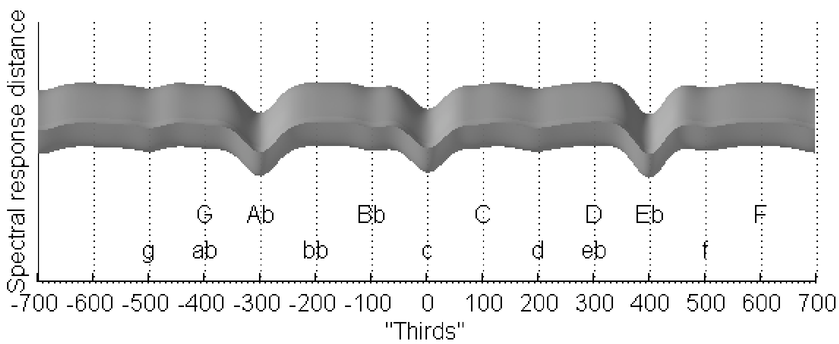
(d)



(e)



(f)



tinuum triads and the reference triad. The greyscale indicates the spectral response distance (the lighter the colour the greater the spectral response distance). The remaining figures view (a) and (d) from the side, so their vertical axes show the spectral response distance; in (b) and (e) the horizontal axes are equivalent to the horizontal axes of (a) and (d); in (c) and (f), the horizontal axes are equivalent to the vertical axes of (a) and (d). A selection of continuum triads is labelled.

3.4.2. Charting Tonal Activities

Figure 3.3 uses the same layout as Figure 3.2 (a) and (d) to show the most important activities implied by the model (those due to parallel comparisons). The reference triad is either C (in which case the blue arrows are used) or c (in which case the pink arrows are used). The arrows indicate the direction of activity between the continuum triad comparison pairs—each arrow points from the triad with negative activity to the comparison triad with positive activity. So, given a reference triad of C, a blue arrow points from a continuum triad to another continuum triad that has a higher spectral response distance and so is heard as its alteration; given a reference triad of c, a pink arrow points from a continuum triad to another continuum triad that is heard as its alteration.

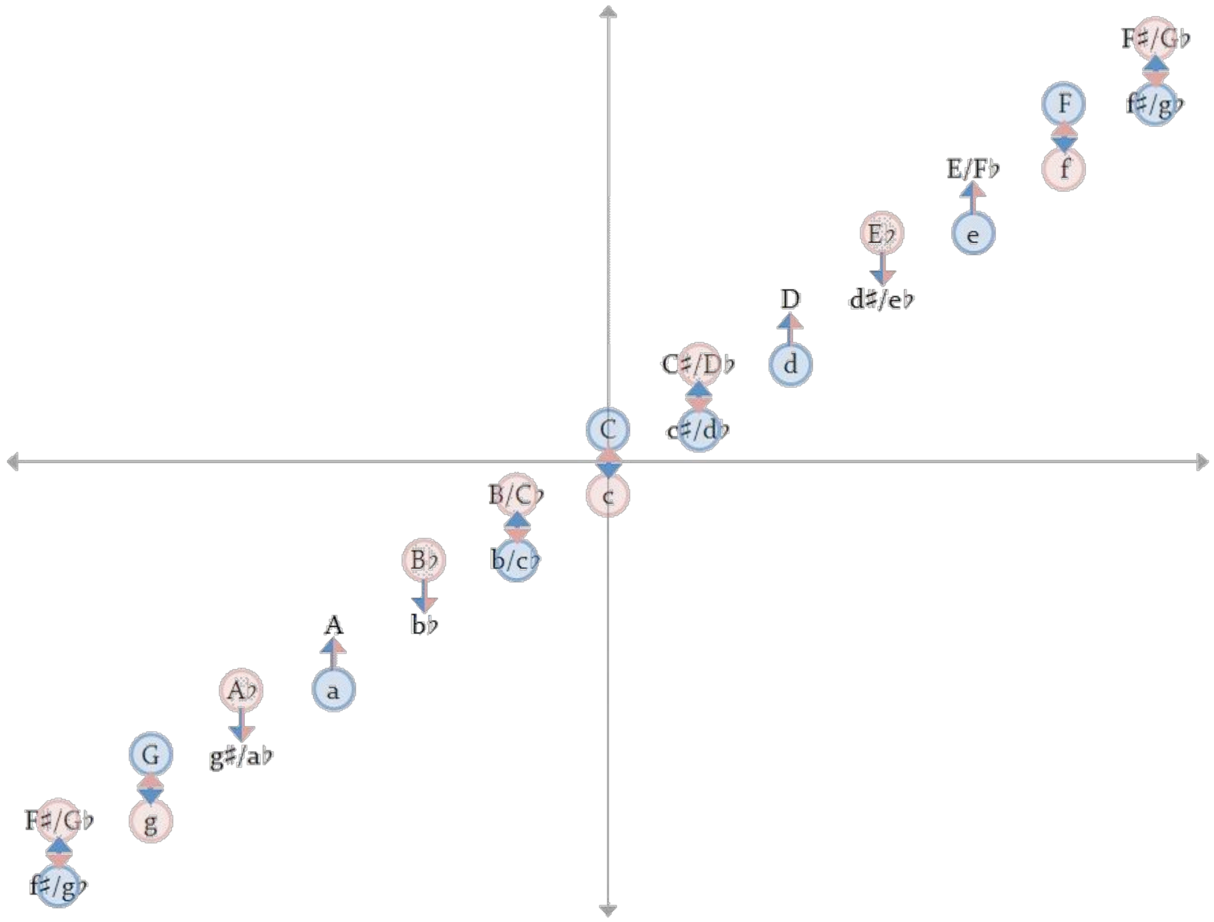


Figure 3.3. A chart of the tonal activities of all 12-TET triads in reference to a C or c triad. The vertical arrows represent activities due to **P** comparisons. Prototypical continuum triads are circled.

When a pairing of a continuum triad W and the reference triad C is inactive in both directions (i.e., $\text{act}_c(\text{C} \rightarrow W|\mathbf{P}) < 0$ AND $\text{act}_c(\text{C} \leftarrow W|\mathbf{P}) < 0$, it is circled in blue; when a pairing of a continuum triad T and the reference triad c is inactive in both directions (i.e., $\text{act}_c(\text{c} \rightarrow W|\mathbf{P}) < 0$ AND $\text{act}_c(\text{c} \leftarrow W|\mathbf{P}) < 0$), it is circled in pink. Such pairings might be considered to be somewhat *prototypical* in that they invoke no activity, and any other pairing is likely to be heard as an alteration of one these prototypes. Indicating the prototypical pairings in this way also enables double alterations to be easily noticed—when proceeding from C to any triad that is not circled and has an arrow pointing towards it from a triad that is circled in pink, such a triad pair is a double alteration; similarly, when proceeding from c to any triad that is not circled and has an arrow pointing towards it from a triad that is circled in blue, such a triad pair is a double alteration. See, for example, the double alteration pairs: c↔D, C↔e♭, and c↔E.

In the following section (which aims to explain familiar cadential progressions, and to shed light upon a number of other regularities of tonal-harmonic music), it may be helpful to refer back to this chart to get an effective visualisation of the analyses undertaken.

4. Non-Experimental Testing of the Model

The principal test of any model or theory is to measure its predictive power—the scope of its predictions and their accuracy. The accuracy of predictions can be assessed by comparing them with either surveyed data, or experimentally generated data. The former has the advantage of ecological validity, but the disadvantage of an inability to isolate the impact of different variables, as well as the inevitable errors and approximations required when collecting the data. The latter has the advantage of greater precision, control, and an ability to isolate, but may lack ecological validity.

I will use both approaches to test the model. My main focus is on experimental testing (discussed in Sect. 5), but in this section I discuss how the model correlates with widely accepted regularities of harmonic tonality. Ideally these regularities would be determined by a thorough statistical analysis of a large database of existing music. Such a task would require the writing of special software (as well as access to a large database of symbolically notated music); this is beyond the scope of this thesis, but may become at least one area for future research.

The next best option is to examine those regularities that are common knowledge, or can be readily identified from pedagogical music theory text books (it is reasonable to assume these provide a good reflection of actual practice). The subsections below provide a survey of these regularities. I cannot claim the survey to be comprehensive, and my own choices will almost certainly show some bias, so any correlations between the regularities and the model must be viewed with some caution. However, this approach is important because it allows the model to be tested against a broader scope of enquiry than is possible with an experimental approach, and the data is ecologically valid.

4.1. Cadences

Music theory typically asserts that the most emphatic root-position major and minor triad cadences are the *authentic* cadences IV→V→I, ii→V→I, iv→V→i, and iv→V→I. Less emphatic cadences are the *deceptive* cadences, the most common examples of which are IV→V→vi, and ii→V→vi. Another type of cadence is the *tritone substitution* cadence, which is an important part of later tonal-harmonic music (particularly jazz), and involve one, or more, chords of the above cadences being transposed by a tritone; for example, VII→♭II→I (which is IV of ♯4→V of ♯4→I).

The following subsections examine the cadential implications predicted by the model for a selection of active antepenult→penult progressions (i.e., progressions where $\mathrm{act}_c(P \to F|\mathbf{P})$ is positive); a similar process can be conducted for any active triad pair to determine cadentially effective progressions.

4.1.1. IV→V→ Cadences

Figure 3.3 shows that C→D is heard as an alteration of C→d (i.e., $\mathrm{act}_c(P \to F|\mathbf{P})$ is positive) so the note *f*♯ is heard as an upwards alteration of the note *f*. To most effectively resolve this alteration, we require a final triad that is in a prototypical pairing with both C and D (i.e., $\mathrm{act}(P \to F|\mathbf{P})$, $\mathrm{act}(P \leftarrow F|\mathbf{P})$, $\mathrm{act}(A \to F\,|\mathbf{P})$, and $\mathrm{act}(A \leftarrow F|\mathbf{P})$ are all negative), and that the altered note *f*♯ moves, preferably upwards, to a note in the final triad, preferably the root (the final should not contain the active tone otherwise no resolution has occurred).

Figure 3.3 shows that the only finals in a prototypical pairing with the antepenult C are C, c♯, d, e, F, f♯, G, a, b and—mentally transposing Figure 3.3 from a C reference to a D reference—the only finals in a prototypical pairing with the penult D are D, d♯, e, f♯, G, g♯, A, b, c♯. The only triads common to both groups are c♯, e, f♯, G, and b. Of these, the only triads that do not contain *f*♯ (the active note) are c♯, e, and G. Of these, the root of G provides an upwards resolution for the active tone, the root of e provides a downwards resolution of the active tone, the root of c♯ is not a resolution for the active tone. This suggests the following ranking of cadential effectiveness (from more effective to less effective): C→D→G (IV→V→I); C→D→a (IV→V→vi); C→D→c♯ (IV of ♭5→V of ♭5→i). This corresponds to conventional theory, in that it gives the highest effectiveness to the authentic cadence, the next highest to the deceptive, and the next highest to the tritone substitution of the authentic.

4.1.2. ii→V→ Cadences

The same process can be carried out for cadences with an antepenult of ii and a penult of V, by starting with the chord pairing c→F. Figure 3.3 shows this is likely to be heard as an alteration of c→f, so the note *a*♮ is heard as an upwards alteration of *a*♭. To most effectively resolve this alteration, we require a final triad in a prototypical pairing with both c and F, and that the altered note *a*♮ moves, preferably upwards, to a note in the final triad, preferably the root.

The only triads that are in a prototypical pairing with both c and F are B♭ and g; of these the root of B♭ provides an upwards resolution for *a*♮, while the root of g provides a downwards resolution. This suggests two cadences, in order of decreasing effectiveness, c→F→B♭ (ii→V→I) and c→F→g (ii→V→vi), which again accords well with conventional music theory—the first being a commonly used authentic cadence, the latter a commonly used deceptive cadence.

4.1.3. iv→V→ Cadences

Figure 3.3 shows that c→D is heard as a double alteration of C→d, and so has two active tones: *f*♯ (which is heard as an upwards alteration of the *f*) and *e*♭ (which is heard as a downwards alteration of *e*). Of these two tones, *f*♯ is the most salient because it is in the penult (and so more recently heard). To most effectively resolve these alterations we require a final triad that is in a prototypical relationship with both the antepenult and penult, and that both altered notes can move, preferably in the same direction as their alteration, to a tone of the final, one of which is, preferably, the root.

For this antepenult→penult progression, there is no final that is prototypical in reference to both. Relative to c, the following triads are prototypical: c, D♭, E♭, f, F♯, g, A♭, B♭, C♭. Relative to D, the following triads are prototypical: D, d♯, e, f♯, G, g♯, A, b, c♯. Of these, D and A make doubly active pairings with c, while c and f make doubly active pairings with D, which leaves the singly altered e, G, g, and B♭ as the best available choices (in terms of activity) for a final. The roots of G and g provide an upwards resolution for the most salient altered tone (*f*♯), the root of e provides a downwards resolution for *f*♯, and the root of B♭ is not a resolution for the active tone. This suggests the following ranking of cadential effectiveness (from more effective to less effective): c→D→G or g (iv→V→I or i); c→D→e (iv→V→vi); c→D→B♭ (iv of ♯4→V of ♯4→I) (note that this last tritone substitution cadence is more effective when the antepenult and penult are reversed to give D→c→B♭ (V of ♯4→iv of ♯4→I), because now the salient active note is the downwards alteration *e*♭, which can resolve downwards to the third of the final). Once again this accords well with conventional music theory, with the authentic cadence ranked highest, then the deceptive, and then the tritone substitution.

4.1.4. ♭II→V→ Cadences

C→F♯ is heard as an alteration of C→f♯, so the note *a*♯ is heard as an upwards alteration of *a*♮. To most effectively resolve this alteration, we require a final triad in a prototypical pairing with both C and F♯, and that the altered note *a*♯ moves, preferably upwards, to a note in the final triad, preferably the root (the final should not contain the active tone otherwise no resolution has occurred).

There is no final that is prototypical in reference to both C and F♯. Relative to C, the following triads are prototypical: C, c♯, d, e, F, f♯, G, a, b. Relative to F♯, the following triads are prototypical: F♯, g, g♯, a♯, B, c, C♯, d♯, e♯. Of these, g♯, a♯ and d♯ make doubly active pairings with C, while d, a, and e make doubly active pairings with F♯, which leaves the singly altered C, c, C♯, c♯, F, e♯, F♯, f♯, G, g, B, b as the best available choices (in terms of activity) for a final. Of these, the only triads that do not contain the active tone *a*♯ are C, c, C♯, c♯, F, e♯, f♯, G, B, b. The roots of B and b provide an upwards resolution for *a*♯, and so make the most cadentially effective final—giving the cadences C→F♯→B (♭II→V→I) and C→F♯→b (♭II→V→i). These are familiar cadences using the Neapolitan ♭II triad (though this triad is typically played in first inversion, in order to improve the voice-leading). The remaining finals provide cadences with varying degrees of effectiveness—many of the embedding scales contain three consecutive semitones and likely have contradictory latent activities. Furthermore, it is likely that the **S** + **P** comparison is important for the C→F♯ progression because C↔F and C↔G have substantially lower spectral response distances than C↔F♯.

4.2. Tonal Functionality

One of the central tenets of functional theory is that all chords can be categorised as belonging to one of three different functions: subdominant, dominant and tonic. When considering only major and minor triads, subdominant is typically represented by IV, iv, ii, or ♭II (these triads are all on the flat, or subdominant, end of the chain of fifths), dominant is typically represented by V (which is at the sharp, or dominant, end of the chain of fifths), and tonic by I and vi (which are in the middle of the chain of fifths).

For triadic harmony, the theory presented here has a similar structure. It requires three triads (antepenult→penult→final) and, as discussed in the previous subsection, effective cadences typically have an antepenult of ii, IV, or iv (i.e., all subdominant function triads), a penult of V (the archetypal dominant function triad), and a final of I or vi (i.e., both tonic function triads).

4.3. Tonal Asymmetry

Functional music theory, and general descriptions of musical practice (see, e.g., Piston & Devoto (1987), and Dahlhaus (1990)), indicate that cadences generally follow the pattern subdominant→dominant→tonic. Reversing the order of the subdominant and dominant chords is generally considered to significantly weaken the effectiveness of the cadence.

This is an example of a tonal asymmetry—the order of presentation of the harmony significantly impacts upon the effect it induces in a listener and, as a result, certain permutations of chords are preferred (and used more often) than others. Other examples of tonal asymmetry can be found in the movement between keys rather than between chords, for example Toiviainen and Krumhansl's (2003) data show that moving from a key to a sharper key is perceived as having greater distance than moving in the opposite direction; similar results were found by Cuddy and Thomson (1991).

As explained in Section 3.2.3.3, act($X \rightarrow Y|\mathbf{P}$) and act($X \leftarrow Y|\mathbf{P}$) are not, in general, equal, and they are frequently of opposite sign. For example, in the pairing C↔D, only D is heard as a parallel alteration, not C. For that reason, the cadence C→D→G (IV→V→I) is likely to be more effective than the cadence D→C→G (V→IV→I) because, in the former, the active tone has extra salience due to it being in the most recently played triad.

Purely structural models, such as the tonal toroid proposed by many contemporary researchers (e.g., see volume 15 of Tonal Theory in the Digital Age: Computing in Musicology) and Lerdahl's tonal pitch space, are inherently symmetrical and so cannot capture the asymmetries that are an important part of our perception of tonality. In the model presented here, tonal asymmetries are a function of the comparisons used—no extra theory needs to be tacked on to account for them.

4.4. Tonal Dualism

Tonal dualism, a concept introduced by Oettingen and Hauptmann, sees major and minor triads as opposites: the major triad is built upwards from the root with a major third and perfect fifth, the minor triad is built downwards from the fifth with the same two intervals. In Riemann, and some of his contemporary followers, tonal dualism is expanded to cover the functions of the scale degrees and triads within the major and minor keys; the functions are typically considered to be reflections of one another. A prominent example of this is the minor key's ♭6, which resolves downwards to the fifth of the minor tonic, which is the dual of the major key's 7, which resolves upwards to the root of the major tonic; the ♭6 sometimes being considered as characteristic of the minor key as 7 is of the major (Harrison, 1994).

Tonal dualism is graphically represented in the plots of spectral response distance—the plots for the major and minor reference triads (Figure 3.2 (a) and (d), respectively) are 180° rotations of each other. For example, with two major triads a whole tone apart (e.g., C↔D), the upper triad (D), not the lower (C), is heard as altered; conversely, with two minor triads a whole tone apart (e.g., c↔d), the lower triad (c), not the upper (d), is heard as altered. If either progression resolves to a G or g tonic: the D has an upwardly altered *f*♯ (7) that resolves upwards to *g* (1); the c has a downwardly altered *e*♭ (♭6) that resolves downwards to *d* (5).

Also notice how in Figure 3.3, the activities of triad pairings whose roots differ by a perfect fifth or fourth (the most common root progressions in Western music) have opposite directions depending on whether the reference triad is major or minor. For example, C→F and c→f are prototypical, while C→f and c→F are active. In C→f the active tone of the latter triad is a downwards alteration, in c→F the active tone of the latter triad is an upwards alteration. If the former resolves to triad a perfect fourth below (e.g., C→f→C), the penult contains the downward resolving ♭6; if the latter resolves to a tonic a perfect fifth below (e.g., c→F→B♭), the penult contains the upward resolving 7.

4.5. Tonal Scales

The modal system of music, which was prevalent until the end of the 16th century, gave no privileged status to any of the modes of the diatonic scale. Tonal music, on the other hand, privileges the Ionian and Aeolian modes; none of the other modes survived into common practice.

The privileged status of these modes is a natural consequence of the model: Given a diatonic scale, the only triad pairs with positive continuous activity are those containing both members of the tritone (e.g., in the "white note" diatonic scale, activity is present only if one triad contains the note *f* (i.e., the triads d and F) and the other triad contains the note *b* (i.e., the triads G and e)). The only triads that make pairings that are prototypical with (d OR F) AND (G OR e) are C and a. These two triads are, therefore, the natural tonics of the “white note” diatonic scale, hence the privileging of the Ionian (whose tonic is C) and Aeolian (whose tonic is a) modes.

4.6. Tonal Harmony

It is interesting to observe that the birth of tonality in the 17th century coincided with the birth of triadic harmony. Balzano (1980) speculates that the privileging of the Aeolian and Ionian modes and the use of triadic harmony are mutually dependent.

The model presented here has a similar dependency—it is only when pairs of triads, rather than pairs of tones or dyads, are used that the conventional effects of tonality are predicted. This mirrors the historical development, and demonstrates a causal dependency of tonality upon triadic harmony.

4.7. Tonal Robustness

The effects of tonal music are robust over the range of tunings used throughout the common practice period (such as meantones, well-temperaments, just intonation, and 12-TET (Barbour, 1951)). They are also robust over a wide range of instrumental timbres (there isn't a different music theory for each different type of instrument).

The model has, so far, been calculated over a variety of meantone tunings and 12-TET, and the resulting *srd* plots, and the activities they imply, are broadly similar. Furthermore, the model does not (like so many others, such as Lerdahl's (2001), or Woolhouse's (2007)) rest upon an implicit assumption of twelve-tone equal temperament—a point that is crucial given that tonality was born in the 17th century when the most common tuning was quarter-comma meantone and 12-tone equal temperament was no more than a gleam in Apollo's eye.

The predictions are also robust over different timbres. The calculations have been made with a spectrum where the partials have amplitudes of $1/i$, where i is the harmonic number of the partial. But the model produces broadly similar results for amplitudes of $1/i^{-d}$ for a wide range of values of d (e.g., $0 < d < 2$).

4.8. Discussion and Conclusion

The data surveyed in this chapter support the conclusion that the psychoacoustic model can explain a broad range of important regularities found in tonal-harmonic music. A non-experimental survey such as this is not, however, the most stringent of tests; the following section describes and analyses two rigorous experiments designed to test the model against human ratings.

5. Experimental Testing of the Model

In the previous section, I discussed how the model can be used to explain many of the regularities of tonal-harmonic music. In this section, I describe two experiments designed to empirically test the model's effectiveness at predicting the reactions of 35 listeners to prepared examples of music.

5.1. Method

The cognitive variables voice-leading distance, spectral distance, and tonal activity cannot be directly measured. But, with careful experimental design, I hoped it would be possible to get a good indication of their values in response to musical stimuli. In two experiments, 35 participants rated the "similarity" and "fit" of the two triads in 26 different triad pairs, and the "cadential effectiveness" of 72 different three-triad progressions.

I expected "similarity" would correspond to voice-leading distance, and "fit" to spectral distance, though I also expected the participants would somewhat confound the two variables, as well as be influenced by unforeseen factors. Despite this, I hoped the results would still be pure enough to allow the ratings to be successfully regressed on pitch distance, fundamental response distance, spectral response distance, and tonal activity, thus providing a useful test of the psychoacoustic model and its underlying theory. I expected the ratings for "cadential effectiveness" to be accurate enough to test the model by regressing it on the spectral response distances and tonal activities of each triad pair.

5.1.1. Participants

There were 35 participants (19 male, 16 female, with a mean age of approximately 30 years), most of whom were students or staff of Jyväskylä University, Finland. Participants were asked to rate their instrumental and music theory skills. The average level of both was "intermediate", and only two participants had no playing or music theory skills (on a scale of "none" = 0, "basic" = 1, "intermediate" = 2, "advanced" = 3, average instrumental skill was 2.3, average music theory skill was 2.1).

The average level of self-reported "reliability" was 66% (participants were asked to estimate the percentage of stimuli they would give the same, or a similar, answer to if they were to do the test again). Participants were asked if they had done analytical listening (i.e., working out the mode of the chords and the root relationships between them): 16 participants claimed to have done no analytical listening; 13 claimed to have done "some" or "a little" analytical listening; 4 to have mostly listened analytically; and 1 to have listened analytically all the time (this participant claimed to possess absolute pitch).

5.1.2. Apparatus and Procedure

The experimental interface (see Appendix D) was created with Max/MSP. The music was stored as MIDI files and played through a software sampler to emulate a string quartet (the synthesizer was Cakewalk's Dimension Pro playing a sample set from Garriton). A string quartet was chosen because, after discussions with colleagues, it was felt to be more pleasant than listening to a purely synthetic sound, and because it lends itself to the hearing out of four independent melodic parts.

The experiments were conducted back-to-back in a quiet room, and the music was played on headphones, with the individual instruments panned to provide a naturalistic stereo image. The first experiment (to rate the "similarity" and "fit" of pairs of triads) took approximately 10 minutes, the second experiment (to rate the cadential effectiveness of three-triad progressions) took approximately 20 minutes.

5.1.3. Stimuli

For each chord progression, voice-leadings were chosen according to standard rules of harmony: there were four parts; common tones and steps were used rather than leaps; parallel fifths and octaves were completely avoided; hidden fifths and octaves were avoided when possible and, when unavoidable, approached by step in one part (given some of the very unusual triad pairings required, hidden fifths cannot always be avoided without creating unpleasant leaps).

The order of presentation was individually randomised for each participant, the tuning was conventional twelve-tone equal temperament (12-TET), and the precise pitch of every chord progression was randomised (in 12-TET steps) over an octave. In between each progression, a short sequence of randomly generated chords was played to lessen the possibility of the previous progression colouring the response to the next.

5.1.3.1. Triad Pairs—"Similarity" and "Fit"

In the first experiment, every participant was asked to rate all possible pairs of 12-TET triads (when disregarding order and overall transposition, there are just 26 different pairs of 12-TET triads) for their "similarity" and "fit". Each triad pair was played as a loop—going from triad 1 to triad 2 to triad 1 to triad 2, and so on. Each chord had a minim (half-note) length, and the tempo was 100 beats (quarter-notes) per minute.

The ratings were made on two separate 5-point scales marked at the bottom and top with "similar" and "dissimilar", and "good fit" and "bad fit", respectively. A value of 1 was given to a rating of maximal similarity or fit (i.e., minimal distance), and a value of 5 to a rating of minimal similarity or fit (i.e., maximal distance). In the instructions, "similar" chords were defined as being those "you might inadvertently think the same"; "dissimilar" with "their difference is obvious and easy to hear"; "good fit" was likened to a chord transition that was "straightforward", "elegant", "easy"; "bad fit" to "clumsy", "awkward", "difficult".

The aim of the "similar/dissimilar" question was to get a rating for voice-leading distance. The aim of the "good fit/bad fit" question to get a rating of spectral distance. It was expected that there would be some confounding of the two concepts, as well as some confounding with other variables (such as activity). But I hoped the ratings would give some indication of the two types of distance.

5.1.3.2. Triad Triples—"Cadential Effectiveness"

In the second experiment, every participant was asked to rate 72 different three-triad progressions for their "cadential effectiveness". Ignoring transposition, there are 1,152 different order-dependent triples of 12-TET triads, so it is unfeasible (in a single experiment) to obtain ratings for all of them. The specific sample of 72 was chosen in order to test the impact of a single type of comparison—the parallel transform—upon discrete activity (and ultimately upon cadential effectiveness), and also to maximise the statistical power of the test at detecting the impact of the variable $\text{act}_\text{d}(A \rightarrow P|\mathbf{P})$.

The parallel comparison (e.g., comparing the spectral distance of triad pair C↔E with the spectral distance of the triad pair C↔e) was chosen because, as discussed in Section 3.2.3.3, it is likely to have the smallest possible voice-leading distance (I show in Sect. 5.2.3 how the experimental data support this conclusion), and so should maximise the absolute value of the continuous activity produced by the two pairs (see Eq. (3.11)). It is likely, therefore, to be the comparison that produces the most salient alteration. Furthermore, as discussed in Section 4, the implications of the parallel comparison provide an effective explanation for many of the regularities of tonal-harmonic music.

Discrete, rather than continuous, activity was chosen for two reasons: firstly, it is relatively easy to choose a sample of three-triad progressions that can isolate changes in the discrete activity of at least one triad pair in a three-triad progression (see below), whereas with continuous activity this is not possible; secondly, it is possible that when judging cadential effectiveness, activity may be more accurately modelled by the simpler discrete measure (due to the cognitive load and long-term memory requirements of the task).

The statistical power for the variable $\text{act}_\text{d}(A \rightarrow P|\mathbf{P})$ was maximised because it seems likely this is one of the more important activity predictors: an effective cadence should be produced when the penult is heard as active as soon as it is played (i.e., directly after the antepenult) so that the following final triad is able to resolve this activity.

The selection of triad triples was made in the following way. The antepenult was either C major or c minor, this makes two possible one-triad "progressions".

The penult was each of the 24 different triads in 12-TET (i.e., the major and minor chords on each degree of the chromatic scale), making a total of 48 different progressions. According to Equations (3.11) and (3.13), for each pair that is a parallel of another (e.g., c→A♭ compared to c→a♭) one pair will have a negative continuous activity (act_c) value, that is, a discrete activity

(act_d) of zero (e.g., c→A♭), the other a positive act_c value, that is, an act_d value of unity (e.g., c→a♭). Hence the latter should be heard as an alteration of the former.

The final chosen for each parallel pair of antepenult→penult pairs was identical (so c→D and c→d get the same final); the root of this final being the resolution of the active tone of the active penult (I have presumed that the best available resolution is made by a semitone step in the same direction as the alteration—see Sect. 3.3.3); the mode of the final was chosen to ensure that $\mathrm{act_d}(A \to P|\mathbf{P}) = 0$; this gives pairs of progressions such as c→D→g and c→d→g.

This selection method is a way of isolating, as much as possible, changes in the value of the discrete activity $\mathrm{act_d}(A \to P|\mathbf{P})$ from changes in the remaining two forwards discrete activity values, $\mathrm{act_d}(P \to F|\mathbf{P})$ and $\mathrm{act_d}(A \to F|\mathbf{P})$ (the reverse activities $\mathrm{act_d}(A \leftarrow P|\mathbf{P})$, $\mathrm{act_d}(P \leftarrow F|\mathbf{P})$, and $\mathrm{act_d}(A \leftarrow F|\mathbf{P})$ are not controlled) as well as other possibly confounding variables such as the scale degrees of the triads' roots or their modes (major or minor). The selection method, therefore, provides four groups with the following patterns of discrete forward activities. Using the simple shorthand notation of $\mathrm{act_d}(A \to P|\mathbf{P})$ | $\mathrm{act_d}(P \to F|\mathbf{P})$ | $\mathrm{act_d}(A \to F|\mathbf{P})$, the four *groups* are: Group 1 = 1 | 1 | 0, Group 2 = 0 | 1 | 0, Group 3 = 1 | 0 | 0, and Group 4 = 0 | 0 | 0. For every member of Groups 1 and 3, there is a member of Group 2 or 4 that has exactly the same triads (ignoring transposition) except for the penult, which has a different mode. The value of having paired groups is that it helps to reduce the degree to which uncontrolled variables contaminate the experiment. Each group contains essentially the same elements, but with the variable of interest, $\mathrm{act_d}(A \to P|\mathbf{P})$, being changed. Note also that all the progressions have $\mathrm{act_d}(A \to F|\mathbf{P}) = 0$, eliminating the impact of this variable.

When these 48 progressions are transposed to give a final major or minor triad with the same root (e.g., C or c), there are just eight different penult→final endings (G→C, G→c, g→C, g→c, B♭→C, B♭→c, b♭→C, and b♭→c). Of these, only four (G→C, G→c, b♭→C, or b♭→c) have a penult with a discrete activity of unity (i.e., they are members of Groups 1 and 3, which have $\mathrm{act_d}(A \to P|\mathbf{P}) = 1$). The final 24 progressions (which make up the total of 72) use these four different penult→final endings but use all antepenults that give both $\mathrm{act_d}(A \to P|\mathbf{P}) = 0$ and $\mathrm{act_d}(A \to F|\mathbf{P}) = 0$. This provides two more groups: Group 5 = 0 | 1 | 0, and Group 6 = 0 | 0 | 0. For every member of Groups 5 and 6, there is a member of Group 1 or 3 or 5 or 6 that has (ignoring transposition) exactly the same penult, a final that has the same root (but not necessarily the same mode), and an antepenult of the opposite mode. As before, these paired groups help to ensure that changes in the value of $\mathrm{act_d}(A \to P|\mathbf{P})$ are isolated from changes in the remaining discrete activities, as well as other potentially confounding variables. Note also, that all 72 progressions have $\mathrm{act_d}(A \to F|\mathbf{P}) = 0$, thus eliminating this variable from the analysis. Appendix C provides a full listing of the triad triples used, and their group numbers.

These related groups are intended to provide an effective way to estimate the impact of $\mathrm{act_d}(A \to P|\mathbf{P})$ upon cadential effectiveness, but this is by no means the only way to select a manageable, but useful, subset of triad triples; however, it does provide a systematic and unbiased method to select those triples that should effectively test the model.

Each triad triple was played through once in full, but the participant could repeat play after a two-second delay. Each chord had a minim (half-note) length, and the tempo was 80 beats (quarter-notes) per minute.

The rating of cadential effectiveness was made on a 7-point scale marked "cadentially effective" at the top, "cadentially ineffective" at the bottom, and "neutral" in the middle. The instructions gave the following explanation of "cadential effectiveness": "how effectively does the third chord give a feeling of 'closure' or 'finality'? For example: If the progression is 'cadentially effective', the third chord gives a clear and definite sense of closure, and would be an effective and unambiguous ending for a piece of music; if the progression is 'cadentially ineffective', the third chord suggests or implies that another chord, or chords, should follow; if the progression is 'neutral', the third chord may give no feeling of closure, but neither does it imply a need for any more chords to follow."

5.2. Results

5.2.1. Similarity

A correlation matrix for the 35 participants' ratings of all 26 triad pairs was created. One participant had three negative correlations with other participants and a low average correlation level (.136), and so was removed as an outlier. The mean of the inter-participant correlations for the 34 remaining participants was .470 (per participant averages ranging from .275 to .599), with no negative values between any pairs of participants, and a Cronbach's α of .967. The inter-participant correlations are given in Table E.1.

For each triad pair, the ratings of similarity were averaged over the 34 participants to create a variable called *sim*. Table 5.1 shows the correlations between *sim* and a number of predictors: the pitch distance moved by the bass (*bas*), tenor (*ten*), alto (*alt*), and soprano (*sop*) voices; the fundamental response distances between the two triads (*frd*); the average of the continuous forwards and backwards activities ($act_c(X \leftrightarrow Y) = \frac{\mathrm{act}_c(X \rightarrow Y) + \mathrm{act}_c(X \leftarrow Y)}{2}$); the average of the discrete forwards and backwards activities ($act_d(X \leftrightarrow Y) = \frac{\mathrm{act}_d(X \rightarrow Y) + \mathrm{d}(X \leftarrow Y)}{2}$).

Table 5.1. Pearson correlations, and their one-tailed significance, between *sim*, *bas*, *ten*, *alt*, *sop*, *frd*, $act_c(X \leftrightarrow Y)$, and $act_d(X \leftrightarrow Y)$.

	sim	*bas*	*ten*	*alt*	*sop*	*frd*	$act_c(X \leftrightarrow Y)$	$act_d(X \leftrightarrow Y)$
sim	1.000	.504	.757	.388	.557	.916	.686	.486
		.004	.000	.025	.002	.000	.000	.006
bas	.504	1.000	.427	.348	.190	.308	.114	.138
	.004		.015	.041	.176	.063	.290	.250
ten	.757	.427	1.000	.140	.639	.760	.386	.289
	.000	.015		.247	.000	.000	.026	.076
alt	.388	.348	.140	1.000	-.034	.317	.258	.069
	.025	.041	.247		.435	.057	.101	.370
sop	.557	.190	.639	-.034	1.000	.715	.251	.275
	.002	.176	.000	.435		.000	.108	.087
frd	.916	.308	.760	.317	.715	1.000	.595	.429
	.000	.063	.000	.057	.000		.001	.014
$act_c(X \leftrightarrow Y)$	.686	.114	.386	.258	.251	.595	1.000	.802
	.000	.290	.026	.101	.108	.001		.000
$act_d(X \leftrightarrow Y)$	.486	.138	.289	.069	.275	.429	.802	1.000
	.006	.250	.076	.370	.087	.014	.000	

All the predictor variables are positively correlated (as expected) with *sim*, and the correlations are all significant ($p < .05$, one-tailed). Table 5.1 indicates there is some collinearity between the predictors, so a stepwise multiple linear regression of *sim* was performed on all the above predictors. The variables *ten*, *alt*, and *sop*, and $act_d(X \leftrightarrow Y)$ drop out; the three remaining variables, *bas*, *frd*, and $act_c(X \leftrightarrow Y)$, provide a highly significant $R^2 = .934$ ($R^2_{adj} = .925$), $F(3, 22) = 103.120$, $p = .000$. Coefficients and their significance for this regression are summarised in Table 5.2, and a scatter plot is shown in Figure 5.1 (note that the two low-valued items are C↔C, and c↔c (and their transpositions), so it makes sense that these are heard as distinctly more similar than all other possible pairs).

Table 5.2. Regression coefficients and significance for multiple regression of *sim* on *frd*, *bas*, and $act_c(X \leftrightarrow Y)$.

	B	*Std. Error*	*β*	*t*	*p*
(Constant)	1.548	.147		32.497	.000
frd	.123	.013	.688	9.607	.000
bas	.115	.025	.265	4.565	.000
$act_c(X \leftrightarrow Y)$	.017	.005	.247	3.595	.002

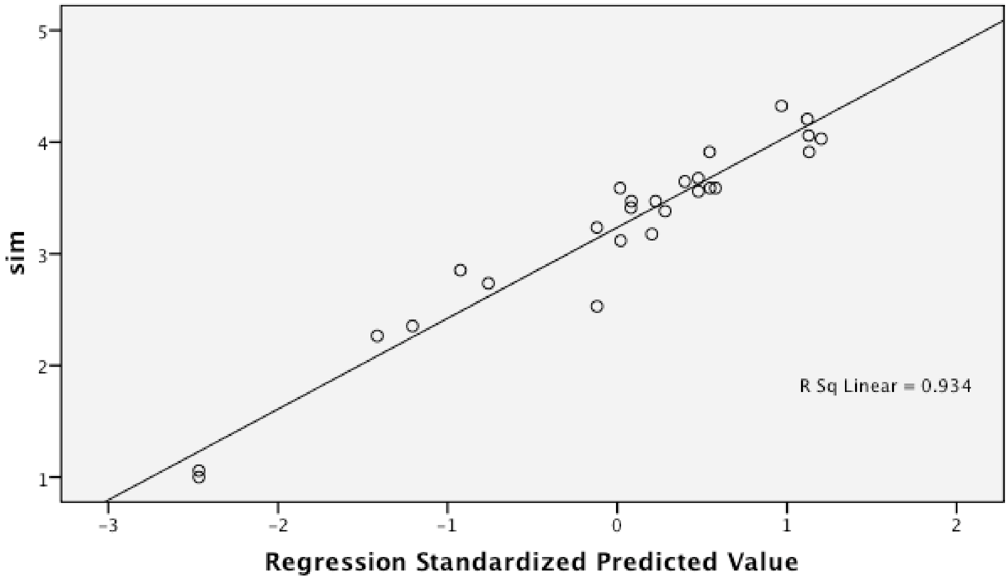


Figure 5.1. Multiple regression of *sim* on *bas*, *frd*, and $act_c(X \leftrightarrow Y)$.

5.2.2. Fit

A correlation matrix for the 35 participants' ratings of all 26 triad pairs was created. Three participants had low average correlation levels (–.048, .004, and .044), and so were removed as outliers. The mean of the inter-participant correlations for the remaining 32 participants was .363

(per participant averages ranging from .165 to .495), with nine negative values between pairs of participants, and a Cronbach's α of .949. The inter-participant correlations are given in Table E.2.

Clearly, the responses for fit were less consistent than those for similarity. Indeed, in the interviews following the test, many participants stated, or implied, that they used familiarity as a strategy—if they recognised a particular progression, they would give it a higher fit. This suggests that these ratings are somewhat affected by each participant's musical taste and familiarity—in other words, a *long-term memory* (*ltm*) component.

For each triad pair, the ratings of similarity were averaged over the 32 participants to create a variable called *fit*. Table 5.3 shows the correlations between *fit*, spectral response distance (*srd*), the average of the continuous forwards and backwards activities ($act_c(X \leftrightarrow Y) = \frac{\mathrm{act}_c(X \to Y) + \mathrm{act}_c(X \leftarrow Y)}{2}$), and the average of the discrete forwards and backwards activities ($act_d(X \leftrightarrow Y) = \frac{\mathrm{act}_d(X \to Y) + \mathrm{d}(X \leftarrow Y)}{2}$).

Table 5.3. Pearson correlations, and their one-tailed significance, between *fit*, *srd*, $act_c(X \leftrightarrow Y)$, and $act_d(X \leftrightarrow Y)$.

	fit	*srd*	$act_c(X \leftrightarrow Y)$	$act_d(X \leftrightarrow Y)$
fit	1.000	.775 .000	.676 .000	.549 .002
srd	.775 .000	1.000	.601 .001	.475 .007
$act_c(X \leftrightarrow Y)$	.676 .000	.601 .001	1.000	.802 .000
$act_d(X \leftrightarrow Y)$	.549 .002	.475 .007	.802 .000	1.000

The predictors $act_c(X \leftrightarrow Y)$ and $act_d(X \leftrightarrow Y)$ are highly collinear, so it makes sense to enter only one of them into a regression of *fit*. The predictor $act_c(X \leftrightarrow Y)$ is better correlated than $act_d(X \leftrightarrow Y)$ with *fit*, so a multiple linear regression of *fit* was performed on *srd* and $act_c(X \leftrightarrow Y)$, giving a highly significant $R^2 = .670$ ($R^2_{adj} = .641$), $F(2, 23) = 23.320$, $p = .000$. Coefficients and their significance for this regression are summarised in Table 5.4, and a scatter plot is shown in Figure 5.2.

Table 5.4. Regression coefficients and significance for multiple regression of *fit* on *srd* and $act_c(X\leftrightarrow Y)$.

	B	*Std. Error*	*β*	*t*	*p*
(Constant)	1.438	.306		4.706	.000
srd	.022	.006	.578	3.853	.001
$act_c(X\leftrightarrow Y)$	.021	.010	.329	2.191	.039

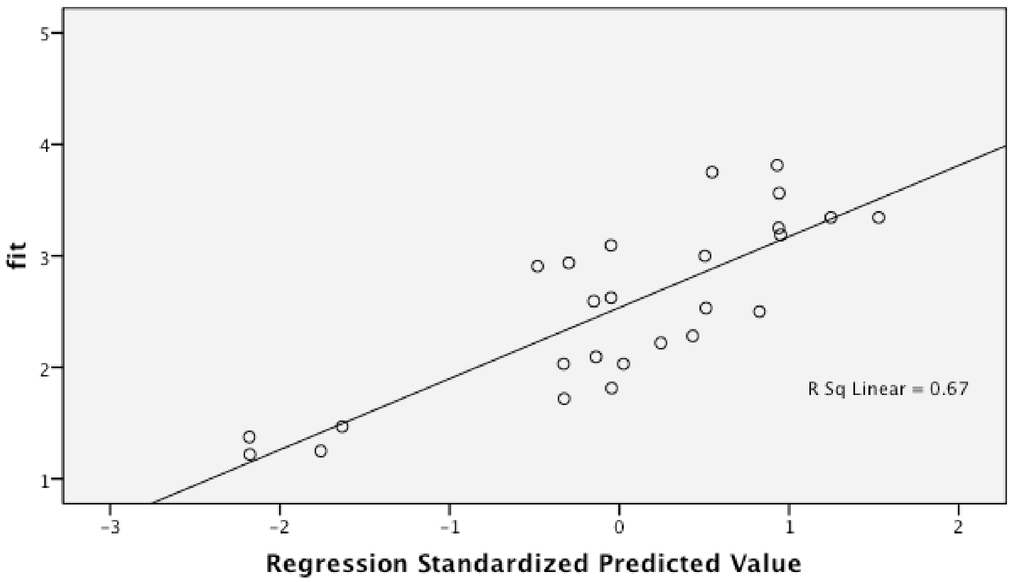


Figure 5.2. Multiple regression of *fit* on *srd* and $act_c(X\leftrightarrow Y)$.

The four low-valued items are C↔C, c↔c, C↔a, and C↔e (and their transpositions), which makes sense from a musical perspective (they are all voice-leading close and diatonic). For this reason, it would be inappropriate to treat them as outliers.

An attempt was made to simulate the effects of the (non-psychoacoustic) long-term memory component by finding all tested cadences that contained a given chord pair. The cadence with the highest-rated cadential effectiveness transferred this rating to an *ltm* rating for that chord pair. The assumption being made here is that if a chord pair is found to be cadentially effective, it is likely to play a prominent and familiar role in music. An example is the progression C↔F♯ (and its transpositions), which was given a much higher rating for *fit* than is predicted from its high spectral response distance and tonal activity. The reason is conjectured to be because it is part of a cadence (the Neapolitan ♭II→V→i) that has a high effectiveness rating and so is commonly used in tonal-harmonic music.

Regressing *fit* with this additional *ltm* variable, significantly increased the regression coefficient (the significance of the change in F was .001) to give $R^2 = .796$ ($R^2_{adj} = .769$), $F(3, 22) = 28.674$, $p = .000$. This suggests not only that fit is influenced by long-term memory, but also that the long-term memory component can be endogenously modelled using calculated values for cadential effectiveness (but that is beyond the scope of this thesis).

5.2.3. Similarity and Fit

The correlations between *sim*, *fit*, *frd*, and *srd*, are shown in Table 5.5.

Table 5.5. Pearson correlations between *sim*, *fit*, *frd*, and *srd*.

	sim	*fit*	*frd*	*srd*
sim	1.000	.741	.916	.907
fit	.741	1.000	.725	.775
frd	.916	.725	1.000	.990
srd	.907	.775	.990	1.000

A Hotelling t-test for the difference between two correlation coefficients from one sample (Hotelling, 1940) shows that the correlation between *frd* and *sim* is not significantly higher than the correlation of *srd* with *sim*, but that the correlation between *srd* and *fit* is significantly higher than the correlation between *frd* and *sim* ($t(23) = 3.047$, $p < .005$, one-tailed). This suggests that *srd* could be substituted for *frd*. However, when the model is developed to enable *srd* to be calculated with different frequency difference limens for different harmonic numbers (see Sect. 3.1.2.1), it is possible that *srd* and *frd* will correlate less. Furthermore, it would be desirable to collect more empirical data to check the correlation of *frd* and *srd* with a variety of different ratings. For this thesis, therefore, it would be premature to explore the implications of such a substitution.

The results of this first experiment suggest the path diagram illustrated in Figure 5.3.

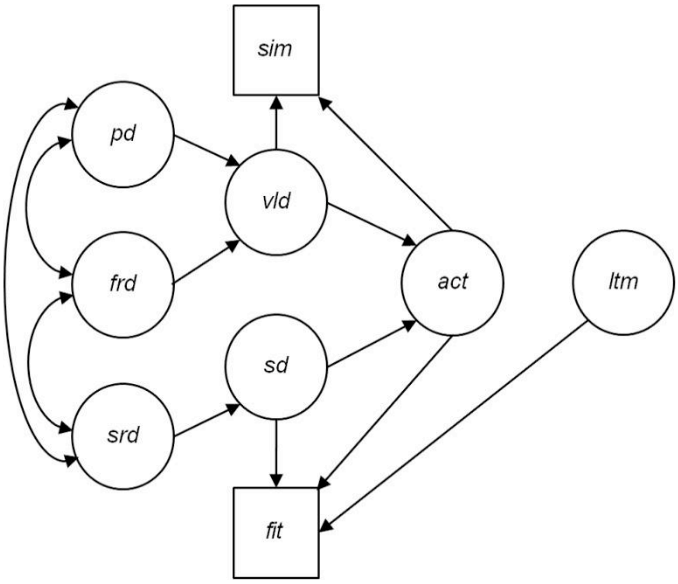


Figure 5.3. A path diagram showing the proposed relationships between the cognitive variables discussed above—including a long-term memory (*ltm*) component—and the measured variables similarity (*sim*) and fit (*fit*). Error terms are not shown.

If the path diagram is correct, these results suggest that not only can *sim* be predicted with great accuracy using *pd*, *frd*, and $act_c(X\leftrightarrow Y)$, but also that the latent voice-leading variable *vld* can be accurately predicted with just *pd* and *frd*. The B values shown in Table 5.2 imply the following formula for voice-leading distance (i.e., they provide values for A_j and B of Eq. (3.7)):

$$\mathrm{vld}(X,Y) = (0.115 \times \mathrm{pd}(\boldsymbol{x}_1,\boldsymbol{y}_1)) + \big(0.123 \times \mathrm{frd}(X,Y)\big), \quad (5.1)$$

for two triads X and Y with bass notes $\boldsymbol{x}_1$ and $\boldsymbol{y}_1$. Using Equation (5.1), Table 5.6 gives the ranking of voice-leading distance for all possible 12-TET triad pairs.

Table 5.6. Example triad pairs are shown in rank order, from smallest to largest, of their *vld* (as calculated with Eq. (5.1)). Common neo-Riemannian abbreviations representing some of the transforms are in the second row.

1	2	3	4	5	6	7	8	9	10	11	12	13
	P	**R**	**L**	**S**								
C↔C	C↔c	C↔a	C↔e	C↔c♯	C↔E♭	C↔E	C↔F	C↔D♭	C↔D	C↔e♭	c↔E	C↔F♯
c↔c					c↔e♭	c↔e	c↔f	c↔D♭	C↔d			C↔f♯
							C↔f	c↔d♭	c↔D			c↔f♯
							C↔f		c↔d			

According to this calculation method, the parallel transform has the smallest non-zero voice-leading distance (which confirms the assumption made to guide the choice of triad progressions for the cadential effectiveness experiment—see Sect. 5.1.3.2).

5.2.4. Cadential Effectiveness

A correlation matrix for the 35 participants' ratings of all 72 triad triples was created. There were no negative correlations between pairs of participants, and the mean of the inter-participant correlations was .494 (per-participant averages ranging from .234 to .621), with a Cronbach's α of .971. The inter-participant correlations are given in Table E.3.

For each triad triple, the ratings of cadential effectiveness were averaged over the 35 participants to create a variable called *eff*. As explained in Section 3.3, the procedure for calculating cadential effectiveness from the psychoacoustic model is multifaceted and complex; ideally it should take into account not just the spectral distance and activity of every chord pairing (in both directions, for the latter), but also trace whether or not each active note is resolved, and whether it resolves to the root, third, or fifth of a triad, as well as undertake analysis of latent progressions in the embedding scale. At the time of writing, these latter factors have not yet been incorporated into the model, but I hope to include them in a future version.

The correlations between *eff*, $act_d(A \rightarrow P|\mathbf{P})$, $act_d(A \leftarrow P|\mathbf{P})$, $act_d(P \rightarrow F|\mathbf{P})$, $act_d(P \leftarrow F|\mathbf{P})$, $act_d(A \leftarrow F|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$ are shown in Table 5.7. Note that $act_d(A \rightarrow F|\mathbf{P})$ is not included because it always has a value of zero for the sample of triad triples tested. Also note that only discrete (act_d), not continuous (act_c), activity values were used because the sample was chosen specifically to test the impact of the discrete activity variable (see Section 5.1.3.2); it can be noted, however, that the continuous activity values are all correlated with *eff* in the expected direction, but with generally lower significance levels than those for discrete activities shown in Table 5.7.

Table 5.7. Pearson correlations, and their one-tailed significance, between *eff*, $act_d(A{\rightarrow}P|\mathbf{P})$, $act_d(A{\leftarrow}P|\mathbf{P})$, $act_d(P{\rightarrow}F|\mathbf{P})$, $act_d(P{\leftarrow}F|\mathbf{P})$, $act_d(A{\leftarrow}F|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$.

	eff	$act_d(A{\rightarrow}P\|\mathbf{P})$	$act_d(A{\leftarrow}P\|\mathbf{P})$	$act_d(P{\rightarrow}F\|\mathbf{P})$	$act_d(P{\leftarrow}F\|\mathbf{P})$	$act_d(A{\leftarrow}F\|\mathbf{P})$	$srd(A, P)$	$srd(P, F)$	$srd(A, F)$
eff	1.000	.205 .042	.036 .382	-.083 .243	-.616 .000	-.024 .420	.207 .040	-.806 .000	-.181 .064
$act_d(A{\rightarrow}P\|\mathbf{P})$	.205 .042	1.000	.000 .500	.393 .000	.250 .017	-.145 .112	.314 .004	.099 .204	-.075 .266
$act_d(A{\leftarrow}P\|\mathbf{P})$	.036 .382	.000 .500	1.000	.225 .029	.000 .500	.415 .000	.273 .010	.042 .364	.088 .231
$act_c(P{\rightarrow}F\|\mathbf{P})$	-.083 .243	.393 .000	.225 .029	1.000	.039 .371	-.006 .481	.353 .001	.311 .004	.023 .424
$act_d(P{\leftarrow}F\|\mathbf{P})$	.393 .000	.250 .017	.000 .500	.039 .371	1.000	.073 .272	.077 .259	.769 .000	.037 .377
$act_d(A{\leftarrow}F\|\mathbf{P})$	-.024 .420	-.145 .112	.415 .000	-.006 .481	.073 .272	1.000	.080 .251	-.017 .444	.256 .015
$srd(A, P)$	.207 .040	.314 .004	.273 .010	.353 .001	.077 .259	.080 .251	1.000	.075 .266	-.198 .048
$srd(P, F)$	-.806 .000	.099 .204	.042 .364	.311 .004	.769 .000	-.017 .444	.075 .266	1.000	-.041 .366
$srd(A, F)$	-.181 .064	-.075 .266	.088 .231	.023 .424	.037 .377	.256 .015	-.198 .048	-.041 .366	1.000

All the predictors show correlations with *eff* in the expected direction (i.e., predictors relating to the final triad have negative correlations, and vice versa); however, not all of the correlations with *eff* are significant. This is probably because the sample set was chosen specifically to maximise the statistical power of the variable $act_d(A{\rightarrow}P|\mathbf{P})$ at the expense of the other activity variables (see Sect. 5.1.3.2) (reassuringly, $act_d(A{\rightarrow}P|\mathbf{P})$ is significantly correlated (p = .042, one-tailed)).

In order to find a more parsimonious model for this sample, a stepwise multiple linear regression was performed on *eff* using all eight variables. The resulting model has four predictor variables: $act_d(A{\rightarrow}P|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$, with R^2 of .794 (R^2_{adj} = .782), $F(4, 67)$ = 64.693, p = .000. Coefficients, and their significance, for this regression are summarised in Table 5.8, and a scatter plot is shown in Figure 5.4.

Table 5.8. Regression coefficients and significance for multiple regression of *eff* on $act_d(A \rightarrow P|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$.

	B	*Std. Error*	*β*	*t*	*p*
(Constant)	9.845	.505		19.495	.000
$act_d(A \rightarrow P\|\mathbf{P})$	.677	.177	.224	3.820	.000
$srd(A, P)$	.013	.005	.168	2.819	.006
$srd(P, F)$	-.111	.007	-.848	-15.207	.000
$srd(A, F)$	-.012	.004	-.166	-2.938	.005

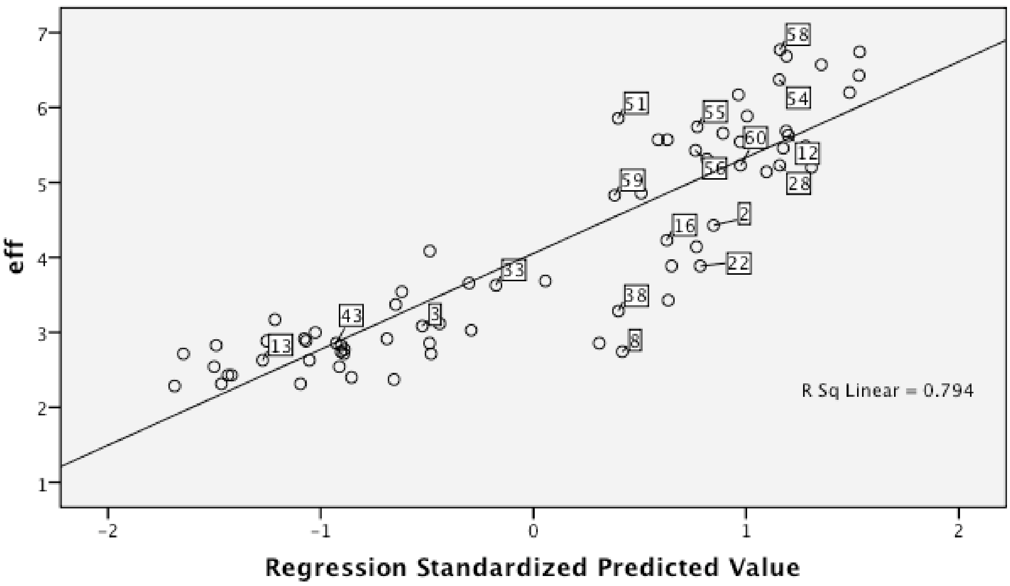


Figure 5.4. Multiple regression of the cadential effectiveness (*eff*) of 72 different triad triples on $act_d(A \rightarrow P|\mathbf{P})$, $srd(A, P)$, $srd(P, F)$, and $srd(A, F)$.

It is interesting to note that most of the act_d predictors (all except $act_d(A \rightarrow P|\mathbf{P})$) drop out in the stepwise regression. This is due, in part, to the high multicollinearities between the act_d variables and their related *srd* variables (e.g., between $act_d(P \rightarrow F|\mathbf{P})$, $act_d(P \leftarrow F|\mathbf{P})$, and $srd(P, F)$). The reason for these multicollinearities is because, when considering all possible voice-leading comparisons, it is probable that any chord pairing with a high *srd* is voice-leading close to another pair with a lower *srd*, so there will be a correlation between *srd* and activity due to all possible comparison pairs. In the current model, only parallel comparisons are considered, but for certain triad pairs (e.g., C↔f♯) other comparisons (e.g., the SLIDE transformation (Lewin, 1987) used to compare C↔f♯ to C↔F) may become important. At this stage, it is impossible to know whether the high *β* values of the three *srd* variables in the regression are due to a direct relationship between *srd* and *eff*, or whether they are only indicative of activity due to comparisons other than just the parallel. It is possible that when a greater variety of triad triples are rated, and more types of comparison are considered, the importance of the *srd* predictors will be diminished and the importance of the activity predictors will be augmented.

It is also interesting to note that many of the "outliers" (both up and down) in this regression are those triad triples that contain only prototypical triad pairs (prototypical pairs have a discrete activity of zero in both directions—see Sect. 3.4.2), such as B♭→g→c and C→G→C (all 18 prototypical-only progressions are labelled in Figure 5.4). Removing all the prototypical-only progressions, and performing the same regression as above, gives an R^2 of .843 (R^2_{adj} = .830), $F(4, 49)$ = 65.619, p = .000. It may be speculated that a progression without active triad pairs lacks cadential "cues" and that the listener may, therefore, be more strongly influenced by long-term memory—that is, does the presented chord progression (which has limited cadential cues) sound like an abbreviation of a familiar progression that does have strong cues that make it cadentially effective, or ineffective? The model predicts that the most effective cadences end with G→C (and its transpositions). This may explain why those prototypical-only progressions that have an ending of G→C (and its transpositions) were typically rated as more cadentially effective than the model predicts, while prototypical-only progressions ending in g→c were typically rated as having a lower cadential effectiveness than the model predicts. This suggests it may be possible to endogenously model the long-term memory component required by the cadential effectiveness model—in a way similar to the *ltm* component implicated in the model for *fit*—but that is beyond the scope of this thesis.

5.3. Discussion and Conclusion

The "similarity" experiment shows that *sim* can be very accurately predicted (R^2 = .934) with *bas*, *frd*, and $act_c(X\leftrightarrow Y)$. This indicates that, when judging the "similarity" of two major or minor root-position triads, a listener is primarily focussed upon their number of common tones (β = .688), then by the pitch distance moved by the bass note (which, for root-position triads, is equivalent to the root note) (β = .265), then by the level of activity due to that pair's parallel comparison (β = .247).

The "fit" experiment shows that *fit* can be reasonably well predicted (R^2 = .670) with *srd* and $act_c(X\leftrightarrow Y)$. This indicates that, when judging the "fit" of two major or minor root-position triads, a listener is primarily focussed upon the number of melodic dyads between them that are spectrally similar (e.g., perfect fifths and fourths, and major and minor thirds and sixths) (β = .578), then by the level of activity due to that pair's parallel comparison (β = .329). It is also likely that long-term memory plays an important role in judgements of "fit", and that triad pairs that occur in effective cadences fit better than might otherwise be expected from their *srd* and $act_c(X\leftrightarrow Y)$ values.

If the latent variable voice leading distance (*vld*) is assumed to be equivalent to "similarity", but with the $act_c(X\leftrightarrow Y)$ variable removed, the parameter values determined by the regression of *sim* in the "similarity" experiment can be used to calculate values for voice-leading distance. These calculations indicate that the parallel transform has the smallest (non-zero) voice-leading distance, followed by the relative transform, then the leading tone exchange, then the SLIDE transform.

The "cadential effectiveness" experiment shows that *eff* is correlated in the expected direction with the *srd* and act_d values for each of the three triad pairs involved (significantly so, for all the *srd* values and some of the activity values). Due to multicollinearity between many of the predictors it is impossible (without collecting further empirical data) to specify a best-possible parsimonious model for all possible triad triples. However, a stepwise regression was used to indicate a possible model, with just four predictors, that provides a good prediction of *eff* (R^2 = .794). The 72 progressions used in the experiment represent only 6.25% of all possible root-position progressions with three triads but, for this sample, the experiment indicates that, when judging the "cadential effectiveness" of progressions with three root-position major and minor triads, a listener's response is strongly determined by the spectral response distance and tonal activities of each pair of triads—an effective cadence is judged to have occurred when triad pairs involving the final are spectrally similar and/or inactive, and triad pairs not involving the final are spectrally dissimilar and/or active. It is not possible to determine the relative importance of the spectral response distance and tonal activity components without further stimuli to reduce their collinearity. Furthermore, without further experiments to provide ratings for a more complete range of three-triad progressions, it is impossible to know how well the model will generalise; despite this, the current results suggest the cadential model holds great promise.

Despite being at an early stage of development, the model provides effective predictions for the perceived "similarity" and "fit" of triads in a pair, and the "cadential effectiveness" of three-triad progressions. I expect that enhancements of the model will further increase its accuracy; these enhancements include: different frequency difference limens for different harmonics; activity values for non-parallel comparisons; ability to track the position, in each triad, of active notes and their resolutions, and weight them accordingly; methods to explore the latent activities of embedding scales; endogenously modelled variables to simulate long-term memory effects; the simulation of short-term memory effects; the ability to handle non-root-position chords, and other more complex chords. Furthermore, the precise details of the causal structure may need modification as more chord progressions are tested. I expect the implementation of these enhancements, as well as collecting ratings for a more complete range of chord progressions, to form part of my future research.

6. Discussion and Conclusion

Both the experimental and non-experimental data support the conclusion that the psychoacoustic model effectively explains how successive triads induce feelings of expectation and resolution (as exemplified by harmonic cadences).

The non-experimental data also support the conclusion that, in addition to cadences, the model explains a broad range of tonal-harmonic regularities; these include tonal asymmetries, functional characteristics, the interdependence of tonality and harmony, and the privileging of certain modes.

Additionally, because this model uses variables based upon empirically derived data and connects them, in a logical fashion, to measures of cognitive distance, it seems reasonable to conclude that its variables, and the causal relationships between them, correspond with real perceptual and cognitive processes.

Acoustics and psychoacoustics have had a long and chequered history in music theory, with theorists from Rameau to Riemann and Schenker grappling, uncomfortably, to make acoustical realities fit with musical realities. And today, a typical view is that although psychoacoustics has some use in explaining harmonic consonance and dissonance it cannot provide an adequate explanation for other aspects of harmonic tonality. I hope the model proposed here demonstrates that psychoacoustics can actually explain many of the core regularities of harmonic tonality.

Furthermore, because the model is psychoacoustically based, it opens up the possibility of exploring the tonal implications of musical systems that use microtonal scales, or tones with non-harmonic partials, or tones with partials that are, to a lesser or greater degree, related to the underlying tuning system (see, e.g., the temperaments, and their related non-diatonic scales, identified by contributors to the Alternate Tunings mailing list and catalogued by Erlich (2006), and the non-harmonic spectra and related scales discussed by Sethares (2004) and Sethares, Milne, Tiedje, Prechtl, and Plamondon (2009)).

Such microtonal scales and non-harmonic timbres also open up interesting avenues for future experimental tests of the model: musical examples could be created with chords and progressions that have never been heard before, thus removing many of the long-term memory effects that can so easily contaminate experiments designed to test psychoacoustical features. I expect such experiments to form an important part of my future research.

I hope it is evident that the novel psychoacoustic approach to tonality presented here holds great promise. Indeed, I hope it may herald a return of psychoacoustics to tonal music theory, as well as act as a launch pad for the exploration of the tonal possibilities opened up by non-standard tunings and spectra.

References

Agawu, V. K. (2003). *Representing African Music: Postcolonial Notes, Queries, Positions.* New York and London: Routledge.

Agmon, E. (1996). Conventional harmonic wisdom and the scope of Schenkerian theory: A reply to John Rothgeb. *Music Theory Online, 2*(3).

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

Barbour, J. M. (1951). *Tuning and temperament: A historical suvey.* East Lansing, Michigan: Michigan State College Press.

Bernstein, D. W. (2002). Nineteenth-century harmonic theory. In *The Cambridge history of Western music theory* (pp. 778–811). Cambridge: Cambridge University Press.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5*, 1–30.

Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception, 5*(3), 219–250.

Caplin, W. (1983). Tonal function and metrical accent: A historical perspective. *Music Theory Spectrum, 5*, 1–14.

Carey, N. (2002). Harmonic experience by W. A. Mathieu. *Music Theory Spectrum, 24*(1), 121–134.

Carey, N., & Clampitt, D. (1989). Aspects of well-formed scales. *Music Theory Spectrum, 11*(2), 187–206.

Chew, E. (2000). *Towards a mathematical model of tonality.* Ph.D. dissertation, MIT, Operations Research Center, Cambridge, MA.

Cuddy, L. L., & Thomson, W. F. (1991). Asymmetry of perceived key movement in chorale sequences: Converging evidence from a probe-tone analysis. *Psychological Research,* (54), 51–59.

Dahlhaus, C. (1980). Tonality. In S. Sadie (Ed.), *The new Grove dictionary of music and musicians* (Vol. 19, pp. 51–55). London: Macmillan.

Dahlhaus, C. (1990). *Studies on the origin of harmonic tonality.* (R. O. Gjerdingen, Trans.) Oxford: Princeton University Press.

Erlich, P. (2006). A middle path between just intonation and the equal temperaments, part 1. *Xenharmonikon*(18), 159–199.

Euler, L. (1739). *Tentamen novae theoriae musicae ex certissismis harmoniae principiis dilucide expositae.* Saint Petersburg: Saint Petersburg Academy.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*(1–2), 103–138.

Harrison, D. (1994). *Harmonic function in chromatic music: A renewed dualist theory and an account of its precedents.* Chicago and London: University of Chicago Press.

Helmholtz, H. L. (1877). *On the sensations of tone.* (A. J. Ellis, Trans.) New York: Dover.

Hindemith, P. (1942). *The craft of musical composition: Book 1* 4th ed.. (A. Mendel, Trans.) New York: Schott.

Hotelling, H. (1940). The selection of variates for use in prediction, with some comments on the general problem of nuisance parameters. *Annals of Mathematical Statistics, 11*, 271–283.

Hyer, B. (2002). Tonality. In *The Cambridge history of Western music theory* (pp. 726–752). Cambridge: Cambridge University Press.

Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory, part I: Consonance of dyads. *The Journal of the Acoustical Society of America, 45*(6), 1451–1459.

Kelley, R. T. (2006). A mathematical model of tonal function. *Annual Meeting of Music Theory Southeast.* Chapel Hill.

Klumpenhouwer, H. (2002). Dualist tonal space and transformation in nineteenth-century musical thought. In T. Christensen (Ed.), *The Cambridge history of Western music theory* (pp. 456–476). Cambridge, UK: Cambridge University Press.

Kopp, D. (2002). *Chromatic transformations in nineteenth-century music.* Cambridge: Cambridge University Press.

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

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception, 17*(4), 481–509.

Lerdahl, F. (2001). *Tonal pitch space.* Oxford: Oxford University Press.

Lester, J. (2002). Rameau and eighteenth-century harmonic theory. In T. Christensen, *The Cambridge history of Western music theory* (pp. 753–777). Cambridge: Cambridge University Press.

Levitin, D. J. (2006). *This is your brain on music: Understanding a human obsession.* New York: Dutton.

Lewin, D. (1987). *Generalized musical intervals and transformations.* New Haven, CT: Yale University Press.

Lowinsky, E. E. (1961). *Tonality and atonality in sixteenth-century music.* Berkeley and Los Angeles: University of California Press.

Mickelsen, W. C. (1977). *Hugo Riemann's theory of harmony.* Lincoln: University of Nebraska Press.

Moore, B. C., Glasberg, B. R., & Shailer, M. J. (1984). Frequency and intensity difference limens for harmonics with complex tones. *The Journal of the Acoustical Society of America, 75*(2), 500–561.

Parncutt, R. (1988). Revision of Terhardt's psychoacoustic model of the root(s) of a musical chord. *Music Perception, 6*(1), 65.

Piston, W., & Devoto, M. (1987). *Harmony* 5th ed.. New York: Norton.

Plomp, R., & Levelt, W. J. (1965). Tonal consonance and critical bandwidth. *The Journal of the Acoustical Society of America, 38,* 548–560.

Quinn, I. (2005). Harmonic function without primary triads. *Annual meeting of Society for Music Theory.* Cambridge.

Schenker, H. (1954). *Harmony.* (O. Jonas, Ed., & E. M. Borgese, Trans.) Chicago: The University of Chicago Press.

Schenker, H. (1987). *Counterpoint.* (J. Rothgeb, & J. Thym, Trans.) New York: Schirmer Books.

Schoenberg, A. (1969). *Structural functions of harmony* 2nd ed.. London: Faber & Faber.

Sethares, W. A. (2004). *Tuning, timbre, spectrum, scale* 2nd ed.. London: Springer-Verlag.

Sethares, W. A., Milne, A. J., Tiedje, S., Prechtl, A., & Plamondon, J. (2009). Spectral tools for Dynamic Tonality and audio morphing. *Computer Music Journal, 33*(2).

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of music: A self-organizing approach. *Psychological Review,* (107), 885–913.

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception, 32,* 741–766.

Tymoczko, D. (2004). Scale networks and Debussy. *Journal of Music Theory, 48*(2), 219–294.

Woolhouse, M. (2007). *Interval cycles and the cognition of pitch attraction in Western tonal-harmonic music.* Ph.D. Dissertation, Cambridge University, Department of Music, Cambridge.

Appendix A: Notational Style

In order to provide musical notation that is unambiguous and precise, I have adopted the following conventions (within quotations, however, the original style of notation has been left unchanged).

Absolute Notation

Lower case italic letters (*a*, *b*, ..., *g*) are used for notes. Upper case Roman letters (A, B, ..., G) are used for major triads, lower case for minor triads (a, b, ..., g). For example:

1. *c* = the note *c*;
2. C = the C major triad;
3. c = the c minor triad;

Relative Notation

Arabic numerals (1, 2, ..., 7) are used for notes. Upper case Roman numerals (I, II, ..., VII) are used for major triads, lower case (i, ii, ..., vii) for minor triads. The number and accidental indicate the position of the note, or chord root in relation to a major scale extending from the overall tonic (i.e., the root of I or i). For example:

1. ♯4 = for example, the note *f*♯ in the key of **C**;
2. ♭III = for example, the major triad E♭ in the key of **c**;

Progressions and Pairings

Directional arrows between pairs of triads, notes, or keys, indicate whether the relationship goes in either direction, or just the one indicated; so C→D indicates a progression from C to D, while C↔D indicates a pairing—a progression from either C to D, or D to C.

Appendix B: Mathematical Proofs and Derivations

Cosine Distance Between Two Response Curves

The internal response curve $\mathrm{ir}_x(f)$, as a function of frequency f, produced by a sine tone of frequency x_f is:

$$\mathrm{ir}_x(f) = \mathrm{e}^{-\frac{(f-x_\mathrm{f})^2}{2\mathrm{DL}(x_\mathrm{f})^2}}, \tag{B.1}$$

where $\mathrm{DL}(x)$ is the frequency difference limen at x. The internal response curve $\mathrm{ir}_x(f)$ is, therefore, a Gaussian function of f centred at x_f with a standard deviation of $\mathrm{DL}(x_\mathrm{f})$.

The cosine distance between two response curves $\mathrm{ir}_x(f)$ and $\mathrm{ir}_y(f)$ is

$$\mathrm{rd}_{\cos}(x,y) \equiv \mathrm{rd}_{\cos}\left(\mathrm{ir}_x(f), \mathrm{ir}_y(f)\right) = 1 - \frac{\int_{-\infty}^{\infty} \mathrm{ir}_x(f)\mathrm{ir}_y(f)\mathrm{d}f}{\sqrt{\int_{-\infty}^{\infty} \mathrm{ir}_x(f)^2\mathrm{d}f \int_{-\infty}^{\infty} \mathrm{ir}_y(f)^2\,\mathrm{d}f}}. \tag{B.2}$$

The product of two Gaussians is another Gaussian:

$$\mathrm{e}^{-b_1(u-c_1)^2}\mathrm{e}^{-b_2(u-c_2)^2} = \mathrm{e}^{-\frac{b_1b_2}{b_1+b_2}(c_1-c_2)^2}\mathrm{e}^{-(b_1+b_2)\left(u-\frac{b_1c_1+b_2c_2}{b_1+b_2}\right)^2}, \tag{B.3}$$

and the definite integral, from $-\infty$ to ∞, of a Gaussian is given by

$$\int_{-\infty}^{\infty} \alpha\mathrm{e}^{-\frac{(u+\beta)^2}{\gamma^2}}\mathrm{d}u = \alpha|\gamma|\sqrt{\pi}. \tag{B.4}$$

Equations (B.3) and (B.4) imply that

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\frac{b_1b_2}{b_1+b_2}(c_1-c_2)^2}\mathrm{e}^{-(b_1+b_2)\left(u-\frac{b_1c_1+b_2c_2}{b_1+b_2}\right)^2}\mathrm{d}u = \mathrm{e}^{-\frac{b_1b_2}{b_1+b_2}(c_1-c_2)^2}\left|\frac{1}{\sqrt{b_1+b_2}}\right|\sqrt{\pi}. \tag{B.5}$$

Substituting f for u; $\frac{1}{2\mathrm{DL}(f_x)^2}$ for b_1; $\frac{1}{2\mathrm{DL}(f_y)^2}$ for b_2; x_f for c_1; and y_f for c_2 (see Eq. (B.1)) into Equation (B.5) gives

$$\int_{-\infty}^{\infty} \mathrm{ir}_x(f)\mathrm{ir}_y(f)\mathrm{d}f = \frac{\sqrt{2\pi}\mathrm{e}^{-\frac{(x_\mathrm{f}-y_\mathrm{f})^2}{2(\mathrm{DL}(x_\mathrm{f})^2+\mathrm{DL}(y_\mathrm{f})^2)}}}{\sqrt{\left|\frac{\mathrm{DL}(x_\mathrm{f})^2+\mathrm{DL}(y_\mathrm{f})^2}{\mathrm{DL}(x_\mathrm{f})^2\mathrm{DL}(y_\mathrm{f})^2}\right|}}, \tag{B.6}$$

$$\int_{-\infty}^{\infty} \mathrm{ir}_x(f)^2\mathrm{d}f = \mathrm{DL}(x_\mathrm{f})\sqrt{\pi}, \tag{B.7}$$

and

$$\int_{-\infty}^{\infty} \mathrm{ir}_y(f)^2 \mathrm{d}f = \mathrm{DL}(y_\mathrm{f})\sqrt{\pi}. \tag{B.8}$$

Substituting these into Equation (B.2) gives

$$\mathrm{rd}_{\cos}(\mathrm{x},\mathrm{y}) \equiv \mathrm{rd}_{\cos}\left(\mathrm{r}_x(f),\mathrm{r}_y(f)\right) = 1 - \frac{\sqrt{2}\mathrm{e}^{-\frac{(x_\mathrm{f}-y_\mathrm{f})^2}{2(\mathrm{DL}(x_\mathrm{f})^2+\mathrm{DL}(y_\mathrm{f})^2)}}}{\sqrt{\left|\frac{\mathrm{DL}(x_\mathrm{f})^2+\mathrm{DL}(y_\mathrm{f})^2}{\mathrm{DL}(x_\mathrm{f})^2\,\mathrm{DL}(y_\mathrm{f})^2}\right| \left|\mathrm{DL}(x_\mathrm{f})^2\,\mathrm{DL}(y_\mathrm{f})^2\right|}} \blacksquare \tag{B.9}$$

Approximate Spectral Distance Between Two Tones with Harmonic Partials

A harmonic tone $\boldsymbol{w}$ consists of partials of frequencies: $w, 2w, 3w, \dots, iw, \dots, \infty w$, where $i \in \mathbb{Z}$ is the harmonic number, and $w \in \mathbb{R}$ is the frequency of the fundamental. If a second tone $\boldsymbol{x} = s\boldsymbol{w}$, where $s \in \mathbb{Z}$, then every partial i of $\boldsymbol{x}$ will be at the same frequency as partial si of $\boldsymbol{w}$; that is, $ix = siw$. If a third tone $\boldsymbol{y} = t\boldsymbol{w}$, where $t \in \mathbb{Z}$, then every partial i of $\boldsymbol{y}$ will be at the same frequency as partial ti of $\boldsymbol{w}$; that is, $iy = tiw$.

Therefore, $iw = ix/s = iy/t$, so

$$ix = \frac{isy}{t}, \tag{B.10}$$

and

$$iy = \frac{itx}{s}. \tag{B.11}$$

However, partials isy/t are only harmonics of $\boldsymbol{y}$ if $s/t \in \mathbb{Z}$. Since $s \in \mathbb{Z}$ and $t \in \mathbb{Z}$, Equation (B.10) can be simplified accordingly:

$$ix = isy. \tag{B.12}$$

Similarly, Equation (B.11) can be simplified accordingly:

$$iy = itx. \tag{B.13}$$

This means that, given two harmonic tones $\boldsymbol{x}$ and $\boldsymbol{y}$ with a frequency ratio of s/t (where s and t are coprime), $1/t$ of $\boldsymbol{x}$'s partials match $\boldsymbol{y}$'s partials, and $1/s$ of $\boldsymbol{y}$'s partials match $\boldsymbol{x}$'s.

Let the total approximate spectral response distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, denoted $\sim\mathrm{srd}(\boldsymbol{x},\boldsymbol{y})$, when $s/t = 1/1$ (so all their partials match), equal zero; let the total approximate spectral response distance of $\boldsymbol{x}$ and $\boldsymbol{y}$, when $s/t \to \infty/\infty$ (so none of their partials match), equal unity. Furthermore, if $1/p$ of all the partials from both tones are matched with another partial, let the approximate spectral response distance equal $1 - 1/p$. This relationship is captured by Equation (B.14)—as s and $t \to 1$, $\sim\mathrm{srd}(\boldsymbol{x}, \boldsymbol{y}) \to 0$; as s and $t \to \infty$, $\sim\mathrm{srd}(\boldsymbol{x}, \boldsymbol{y}) \to 1$:

$$\sim\mathrm{srd}(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{1}{2s} - \frac{1}{2t} \blacksquare \tag{B.14}$$

APPENDIX C: TRIAD PAIRS AND TRIPLES

Table C.1. The triad pairs used in the experiment to get ratings for “similarity” and “fit”. The pitch of each pair was randomised over 12 equally tempered semitones. The spectral response distances (*srd*) calculated by the model are also shown.

No.	Pair	*srd*
1	C↔F♯	73.04
2	C↔f♯	69.14
3	C↔G	43.30
4	C↔g	46.92
5	C↔A♭	45.16
6	C↔g♯	65.88
7	C↔A	47.66
8	C↔a	22.64
9	C↔B♭	64.68
10	C↔b♭	69.16
11	C↔B	67.86
12	C↔b	64.89
13	C↔C	4.62
14	C↔c	23.33
15	C↔c♯	51.35
16	C↔d	60.64
17	C↔e♭	71.29
18	C↔e	23.26
19	C↔f	45.89
20	c↔g♭	72.91
21	c↔g	43.28
22	c↔a♭	44.95
23	c↔a	47.63
24	c↔b♭	64.50
25	c↔b	67.70
26	c↔c	4.54

Table C.2. The triad triples used in the experiment to collect ratings for "cadential effectiveness". The pitch of every progression was randomised over 12 equally tempered semitones. The Group number (see Sect. 5.1.3.2) is also shown.

No.	Triple		Group
1	D♭→G→c	♭II→V→i	1
2	D♭→g→c	♭II→v→i	4
3	E♭→B♭→c	♭III→♭VII→i	4
4	E♭→b♭→c	♭III→♭vii→i	3
5	D→B♭→C	II→♭VII→I	2
6	D→b♭→C	II→♭vii→I	1
7	B♭→G→c	♭VII→V→i	1
8	B♭→g→c	♭VII→v→i	4
9	C→B♭→C	I→♭VII→I	2
10	C→b♭→C	I→♭vii→I	1
11	A♭→G→c	♭VI→V→i	1
12	A♭→g→c	♭VI→v→i	4
13	B♭→B♭→c	♭VII→♭VII→i	4
14	B♭→b♭→c	♭VII→♭vii→i	3
15	F♯→G→c	♯IV→V→i	1
16	G♭→g→c	♭V→v→i	4
17	F→G→C	IV→V→I	3
18	F→g→C	IV→v→I	2
19	G→B♭→C	V→♭VII→I	2
20	G→b♭→C	V→♭vii→I	1
21	E♭→G→c	♭III→V→i	1
22	E♭→g→c	♭III→v→i	4
23	F→B♭→C	IV→♭VII→I	2
24	F→b♭→C	IV→♭vii→I	1
25	e→B♭→C	iii→♭VII→I	2
26	e→b♭→C	iii→♭vii→I	1
27	c→G→c	i→V→i	1
28	c→g→c	i→v→i	4
29	d→B♭→C	ii→♭VII→I	2
30	d→b♭→C	ii→♭vii→I	1
31	b♭→G→c	♭vii→V→i	1
32	b♭→g→c	♭vii→v→i	4
33	c→B♭→c	i→♭VII→i	4
34	c→b♭→c	i→♭vii→i	3
35	b→B♭→C	vii→♭VII→I	2
36	b→b♭→C	vii→♭vii→I	1
37	g→G→c	v→V→i	1
38	g→g→c	v→v→i	4
39	a→B♭→C	vi→♭VII→I	2
40	a→b♭→C	vi→♭vii→I	1
41	f→G→c	iv→V→i	1
42	f→g→c	iv→v→i	4
43	g→B♭→c	v→♭VII→i	4
44	g→b♭→c	v→♭vii→i	3
45	e♭→G→c	♭iii→V→i	1
46	e♭→g→c	♭iii→v→i	4
47	d→G→C	ii→V→I	3
48	d→g→C	ii→v→I	2
49	D→G→C	II→V→I	6
50	E→G→C	III→V→I	6
51	G→G→C	V→V→I	6
52	A→G→C	VI→V→I	6
53	B→G→c	VII→V→i	5
54	C→G→C	I→V→I	6
55	e→G→C	iii→V→I	6
56	f♯→G→C	♯iv→V→I	6
57	g♯→G→c	♯v→V→i	5
58	a→G→C	vi→V→I	6
59	b→G→C	vii→V→I	6
60	c♯→G→C	♯i→V→I	6
61	E→b♭→C	III→♭vii→I	5
62	G♭→b♭→c	♭V→♭vii→i	6
63	A♭→b♭→c	♭VI→♭vii→i	6
64	A→b♭→C	VI→♭vii→I	5
65	B→b♭→c	VII→♭vii→i	6
66	D♭→b♭→c	♭II→♭vii→i	6
67	f→b♭→c	iv→♭vii→i	6
68	g♭→b♭→C	♭v→♭vii→I	5
69	a♭→b♭→c	♭vi→♭vii→i	6
70	b♭→b♭→c	♭vii→♭vii→i	6
71	d♭→b♭→C	♭ii→♭vii→I	5
72	e♭→b♭→c	♭iii→♭vii→i	6

Appendix D: Interfaces of the two Experiments

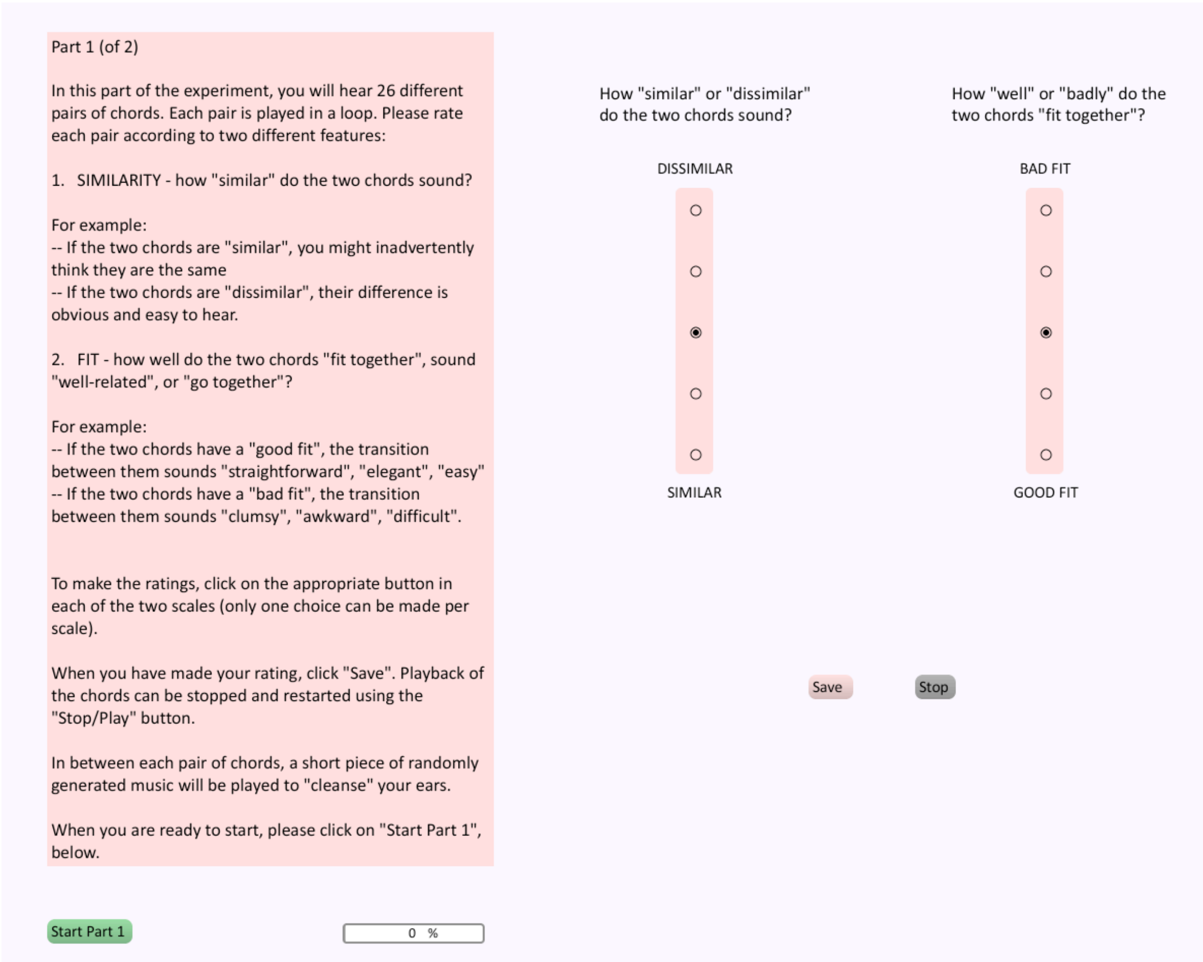


Figure D.1. GUI of the first experiment.

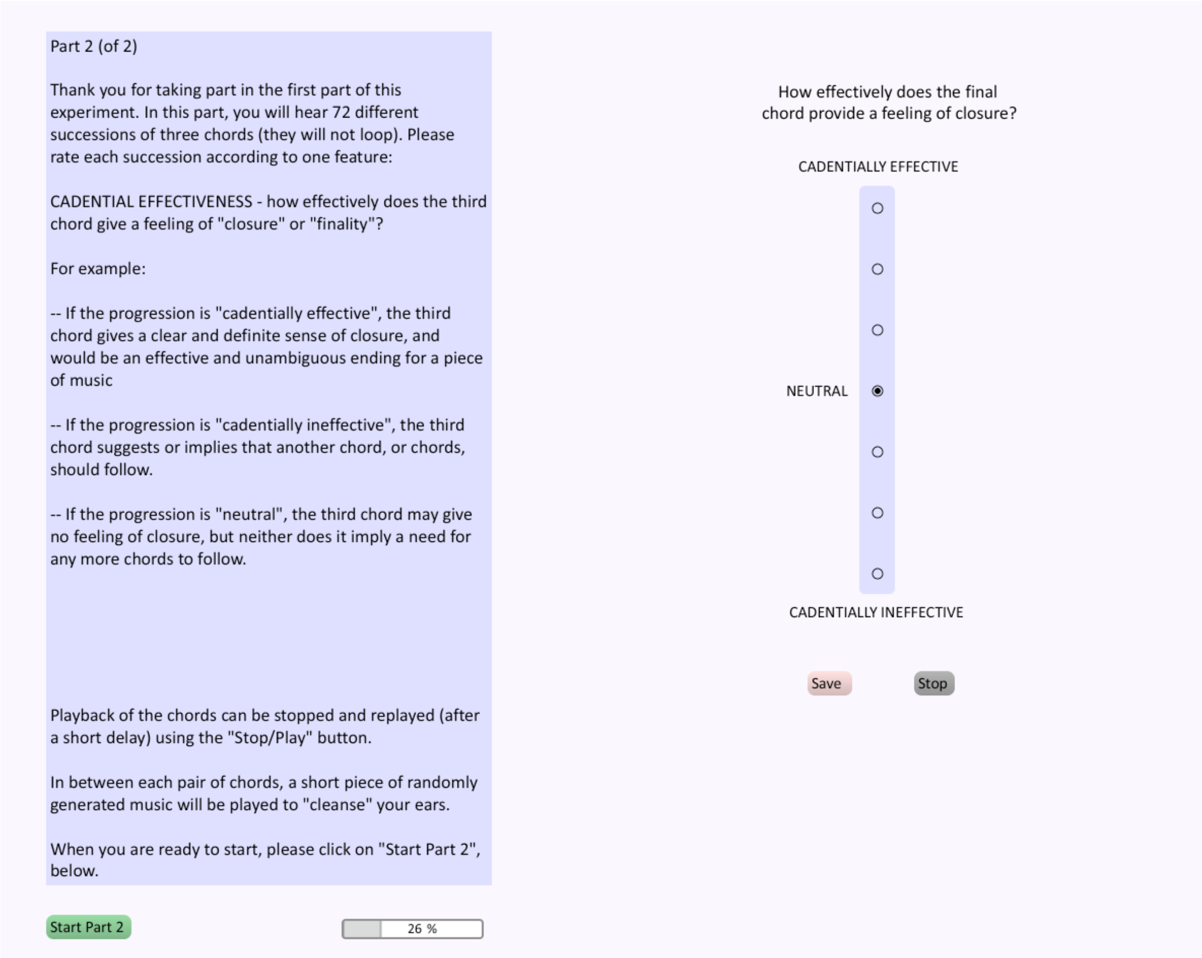


Figure D.2. GUI of the second experiment.

Appendix E: Inter-Participant Correlation Matrices

Table E.1. Inter-participant correlations (Pearson) for ratings of "similarity". The single outlying participant is not shown. The column to the right shows the per participant mean correlation, the value at the bottom right shows the overall mean correlation.

Inter-participant Correlations																																		Mean
1.0	0.6	0.6	0.3	0.6	0.6	0.5	0.4	0.7	0.2	0.6	0.3	0.5	0.4	0.3	0.5	0.6	0.5	0.4	0.4	0.2	0.7	0.6	0.6	0.1	0.6	0.6	0.6	0.5	0.6	0.5	0.4	0.4	0.5	0.477
0.6	1.0	0.6	0.7	0.3	0.6	0.7	0.5	0.7	0.5	0.5	0.7	0.4	0.6	0.5	0.7	0.7	0.7	0.4	0.5	0.3	0.7	0.6	0.6	0.5	0.5	0.7	0.5	0.7	0.3	0.6	0.5	0.3	0.7	0.555
0.6	0.6	1.0	0.5	0.6	0.4	0.6	0.3	0.8	0.3	0.5	0.4	0.6	0.5	0.2	0.7	0.5	0.5	0.5	0.5	0.3	0.7	0.4	0.7	0.2	0.6	0.7	0.7	0.4	0.4	0.5	0.4	0.6	0.4	0.503
0.3	0.7	0.5	1.0	0.3	0.4	0.5	0.4	0.6	0.4	0.5	0.5	0.5	0.6	0.5	0.7	0.6	0.4	0.4	0.5	0.4	0.5	0.5	0.6	0.4	0.5	0.6	0.5	0.4	0.4	0.7	0.3	0.2	0.6	0.482
0.6	0.3	0.6	0.3	1.0	0.4	0.5	0.5	0.6	0.2	0.5	0.2	0.5	0.4	0.4	0.4	0.5	0.4	0.4	0.3	0.4	0.7	0.4	0.6	0.1	0.6	0.5	0.5	0.1	0.4	0.6	0.6	0.6	0.4	0.448
0.6	0.6	0.4	0.4	0.4	1.0	0.4	0.3	0.5	0.4	0.6	0.4	0.4	0.6	0.4	0.5	0.5	0.6	0.2	0.4	0.4	0.5	0.6	0.5	0.2	0.6	0.6	0.6	0.3	0.3	0.5	0.6	0.4	0.6	0.468
0.5	0.7	0.6	0.5	0.5	0.4	1.0	0.3	0.5	0.4	0.5	0.5	0.6	0.4	0.5	0.4	0.6	0.6	0.4	0.7	0.2	0.7	0.7	0.5	0.3	0.4	0.7	0.4	0.3	0.3	0.4	0.5	0.6	0.5	0.489
0.4	0.5	0.3	0.4	0.5	0.3	0.3	1.0	0.5	0.4	0.5	0.5	0.3	0.4	0.2	0.3	0.4	0.4	0.0	0.3	0.4	0.5	0.5	0.4	0.1	0.4	0.2	0.3	0.2	0.4	0.4	0.5	0.2	0.5	0.358
0.7	0.7	0.8	0.6	0.6	0.5	0.5	0.5	1.0	0.3	0.7	0.6	0.6	0.5	0.3	0.7	0.7	0.5	0.4	0.6	0.3	0.8	0.5	0.8	0.3	0.7	0.7	0.5	0.5	0.5	0.6	0.5	0.6	0.5	0.562
0.2	0.5	0.3	0.4	0.2	0.4	0.4	0.4	0.3	1.0	0.3	0.3	0.1	0.6	0.4	0.2	0.4	0.4	0.2	0.4	0.5	0.3	0.4	0.3	0.0	0.2	0.3	0.4	0.4	0.2	0.3	0.6	0.2	0.5	0.332
0.6	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.7	0.3	1.0	0.3	0.4	0.4	0.2	0.5	0.5	0.6	0.2	0.5	0.2	0.7	0.6	0.5	0.1	0.5	0.5	0.5	0.2	0.3	0.5	0.3	0.4	0.4	0.443
0.3	0.7	0.4	0.5	0.2	0.4	0.5	0.5	0.6	0.3	0.3	1.0	0.4	0.6	0.3	0.6	0.7	0.7	0.3	0.7	0.1	0.6	0.4	0.5	0.4	0.5	0.5	0.2	0.6	0.3	0.6	0.5	0.4	0.4	0.448
0.5	0.4	0.6	0.5	0.5	0.4	0.6	0.3	0.6	0.1	0.4	0.4	1.0	0.5	0.4	0.7	0.7	0.5	0.5	0.4	0.3	0.5	0.4	0.6	0.5	0.7	0.6	0.5	0.2	0.4	0.6	0.4	0.5	0.2	0.461
0.4	0.6	0.5	0.6	0.4	0.6	0.4	0.4	0.5	0.6	0.4	0.6	0.5	1.0	0.7	0.6	0.6	0.8	0.5	0.5	0.6	0.6	0.4	0.6	0.3	0.6	0.6	0.6	0.5	0.5	0.7	0.7	0.6	0.6	0.548
0.3	0.5	0.2	0.5	0.4	0.4	0.5	0.2	0.3	0.4	0.2	0.3	0.4	0.7	1.0	0.3	0.5	0.6	0.5	0.5	0.2	0.4	0.4	0.4	0.3	0.5	0.4	0.3	0.3	0.3	0.5	0.6	0.4	0.6	0.408
0.5	0.7	0.7	0.7	0.4	0.5	0.4	0.3	0.7	0.2	0.5	0.6	0.7	0.6	0.3	1.0	0.7	0.5	0.4	0.5	0.4	0.6	0.3	0.7	0.3	0.6	0.6	0.6	0.6	0.4	0.7	0.4	0.3	0.4	0.521
0.6	0.7	0.5	0.6	0.5	0.5	0.6	0.4	0.7	0.4	0.5	0.7	0.7	0.6	0.5	0.7	1.0	0.6	0.5	0.6	0.3	0.7	0.5	0.8	0.5	0.8	0.7	0.5	0.6	0.5	0.8	0.6	0.4	0.5	0.582
0.5	0.7	0.5	0.4	0.4	0.6	0.6	0.4	0.5	0.4	0.6	0.7	0.5	0.8	0.6	0.5	0.6	1.0	0.5	0.6	0.4	0.8	0.5	0.5	0.2	0.5	0.7	0.5	0.4	0.5	0.6	0.7	0.7	0.6	0.544
0.4	0.4	0.5	0.4	0.4	0.2	0.4	0.0	0.4	0.2	0.2	0.3	0.5	0.5	0.5	0.4	0.5	0.5	1.0	0.3	0.3	0.5	0.2	0.5	0.3	0.6	0.5	0.5	0.4	0.3	0.6	0.5	0.4	0.4	0.396
0.4	0.5	0.5	0.5	0.3	0.4	0.7	0.3	0.6	0.4	0.5	0.7	0.4	0.5	0.5	0.5	0.6	0.6	0.3	1.0	0.2	0.7	0.4	0.5	0.1	0.4	0.4	0.3	0.4	0.5	0.5	0.4	0.5	0.5	0.451
0.2	0.3	0.3	0.4	0.4	0.4	0.2	0.4	0.3	0.5	0.2	0.1	0.3	0.6	0.2	0.4	0.3	0.4	0.3	0.2	1.0	0.4	0.1	0.5	0.2	0.4	0.3	0.6	0.1	0.4	0.4	0.5	0.4	0.3	0.338
0.7	0.7	0.7	0.5	0.7	0.5	0.7	0.5	0.8	0.3	0.7	0.6	0.5	0.6	0.4	0.6	0.7	0.8	0.5	0.7	0.4	1.0	0.5	0.7	0.3	0.7	0.7	0.5	0.5	0.6	0.7	0.7	0.7	0.6	0.599
0.6	0.6	0.4	0.5	0.4	0.6	0.7	0.5	0.5	0.4	0.6	0.4	0.4	0.4	0.4	0.3	0.5	0.5	0.2	0.4	0.1	0.5	1.0	0.4	0.2	0.3	0.5	0.3	0.3	0.2	0.3	0.6	0.4	0.6	0.420
0.6	0.6	0.7	0.6	0.6	0.5	0.5	0.4	0.8	0.3	0.5	0.5	0.6	0.6	0.4	0.7	0.8	0.5	0.5	0.5	0.5	0.7	0.4	1.0	0.4	0.8	0.6	0.6	0.6	0.5	0.8	0.6	0.5	0.5	0.572
0.1	0.5	0.2	0.4	0.1	0.2	0.3	0.1	0.3	0.0	0.1	0.4	0.5	0.3	0.3	0.3	0.5	0.2	0.3	0.1	0.2	0.3	0.2	0.4	1.0	0.6	0.6	0.2	0.1	0.0	0.3	0.3	0.3	0.2	0.275
0.6	0.5	0.6	0.5	0.6	0.6	0.4	0.4	0.7	0.2	0.5	0.5	0.7	0.6	0.5	0.6	0.8	0.5	0.6	0.4	0.4	0.7	0.3	0.8	0.6	1.0	0.6	0.6	0.4	0.4	0.7	0.6	0.5	0.5	0.540
0.6	0.7	0.7	0.6	0.5	0.6	0.7	0.2	0.7	0.3	0.5	0.5	0.6	0.6	0.4	0.6	0.7	0.7	0.5	0.4	0.3	0.7	0.5	0.6	0.6	0.6	1.0	0.6	0.4	0.3	0.6	0.6	0.6	0.5	0.542
0.6	0.5	0.7	0.5	0.5	0.6	0.4	0.3	0.5	0.4	0.5	0.2	0.5	0.6	0.3	0.6	0.5	0.5	0.5	0.3	0.6	0.5	0.3	0.6	0.2	0.6	0.6	1.0	0.3	0.4	0.6	0.4	0.4	0.5	0.474
0.5	0.7	0.4	0.4	0.1	0.3	0.3	0.2	0.5	0.4	0.2	0.6	0.2	0.5	0.3	0.6	0.6	0.4	0.4	0.4	0.1	0.5	0.3	0.6	0.1	0.4	0.4	0.3	1.0	0.2	0.5	0.4	0.1	0.4	0.375
0.6	0.3	0.4	0.4	0.4	0.3	0.3	0.4	0.5	0.2	0.3	0.3	0.4	0.5	0.3	0.4	0.5	0.5	0.3	0.5	0.4	0.6	0.2	0.5	0.0	0.4	0.3	0.4	0.2	1.0	0.6	0.4	0.5	0.5	0.387
0.5	0.6	0.5	0.7	0.6	0.5	0.4	0.4	0.6	0.3	0.5	0.6	0.6	0.7	0.5	0.7	0.8	0.6	0.6	0.5	0.4	0.7	0.3	0.8	0.3	0.7	0.6	0.6	0.5	0.6	1.0	0.5	0.5	0.6	0.546
0.4	0.5	0.4	0.3	0.6	0.6	0.5	0.5	0.5	0.6	0.3	0.5	0.4	0.7	0.6	0.4	0.6	0.7	0.5	0.4	0.5	0.7	0.6	0.6	0.3	0.6	0.6	0.4	0.4	0.4	0.5	1.0	0.7	0.7	0.517
0.4	0.3	0.6	0.2	0.6	0.4	0.6	0.2	0.6	0.2	0.4	0.4	0.5	0.6	0.4	0.3	0.4	0.7	0.4	0.5	0.4	0.7	0.4	0.5	0.3	0.5	0.6	0.4	0.1	0.5	0.5	0.7	1.0	0.3	0.444
0.5	0.7	0.4	0.6	0.4	0.6	0.5	0.5	0.5	0.5	0.4	0.4	0.2	0.6	0.6	0.4	0.5	0.6	0.4	0.5	0.3	0.6	0.6	0.5	0.2	0.5	0.5	0.5	0.4	0.5	0.6	0.7	0.3	1.0	0.480
																																		0.470

Table E.2. Inter-participant correlations (Pearson) for ratings of "fit". The three outlying participants are not shown. The column to the right shows the per participant mean correlation, the value at the bottom right shows the overall mean correlation.

Inter-participant Correlations																																Mean
1.0	0.5	0.7	0.6	0.5	0.3	0.6	0.7	0.5	0.7	0.6	0.3	0.4	0.6	0.4	0.5	0.2	0.2	0.4	0.1	0.7	0.7	0.6	0.5	0.4	0.2	0.1	0.5	0.4	0.2	0.4	0.4	0.452
0.5	1.0	0.4	0.2	0.2	0.4	0.5	0.5	0.7	0.6	0.5	0.2	0.1	0.6	0.4	0.5	0.6	0.1	0.3	-0.1	0.7	0.4	0.5	0.4	0.5	0.1	0.3	0.3	0.4	0.0	0.5	0.4	0.380
0.5	0.4	1.0	0.3	0.4	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.2	0.5	0.3	0.3	0.1	0.4	0.3	0.2	0.5	0.4	0.5	0.5	0.5	0.3	0.5	0.2	0.6	0.3	0.3	0.3	0.393
0.7	0.2	0.3	1.0	0.5	0.1	0.3	0.5	0.3	0.5	0.3	0.0	0.0	0.3	0.4	0.6	0.2	0.2	0.5	0.3	0.3	0.3	0.4	0.4	0.3	0.0	0.0	0.2	0.5	0.2	0.4	0.5	0.318
0.6	0.2	0.4	0.5	1.0	0.2	0.5	0.7	0.5	0.6	0.6	0.4	0.3	0.5	0.6	0.3	0.2	0.4	0.3	0.4	0.5	0.5	0.4	0.4	0.5	-0.1	0.1	0.6	0.4	0.3	0.5	0.6	0.416
0.5	0.4	0.6	0.1	0.2	1.0	0.2	0.3	0.3	0.1	0.3	0.4	-0.2	0.2	0.1	0.3	0.4	0.3	-0.1	0.3	0.2	0.1	0.2	0.2	0.5	0.1	0.5	0.0	0.5	0.2	0.2	0.3	0.248
0.3	0.5	0.5	0.3	0.5	0.2	1.0	0.6	0.5	0.6	0.4	0.3	0.3	0.6	0.6	0.6	0.5	0.5	0.5	-0.1	0.5	0.6	0.4	0.4	0.5	0.0	0.0	0.4	0.5	0.2	0.3	0.7	0.413
0.6	0.5	0.5	0.5	0.7	0.3	0.6	1.0	0.6	0.6	0.6	0.4	0.5	0.6	0.4	0.3	0.5	0.5	0.4	0.4	0.7	0.6	0.6	0.7	0.5	0.2	0.2	0.6	0.5	0.3	0.5	0.5	0.495
0.7	0.7	0.5	0.3	0.5	0.3	0.5	0.6	1.0	0.7	0.6	0.4	0.3	0.6	0.7	0.4	0.6	0.1	0.4	0.0	0.6	0.5	0.4	0.4	0.6	0.2	0.3	0.6	0.4	0.1	0.4	0.4	0.444
0.5	0.6	0.5	0.5	0.6	0.1	0.6	0.6	0.7	1.0	0.8	0.3	0.4	0.8	0.5	0.4	0.2	0.2	0.5	0.2	0.7	0.7	0.6	0.3	0.5	0.0	0.2	0.8	0.4	0.1	0.5	0.4	0.457
0.7	0.5	0.5	0.3	0.6	0.3	0.4	0.6	0.6	0.8	1.0	0.4	0.4	0.7	0.4	0.3	0.2	0.2	0.2	0.2	0.7	0.8	0.5	0.3	0.5	0.2	0.4	0.7	0.3	0.1	0.6	0.2	0.433
0.6	0.2	0.5	0.0	0.4	0.4	0.3	0.4	0.4	0.3	0.4	1.0	0.1	0.5	0.4	0.3	0.1	0.6	0.2	0.3	0.4	0.2	0.4	0.2	0.5	0.5	0.5	0.2	0.3	0.2	0.3	0.3	0.338
0.3	0.1	0.2	0.0	0.3	-0.2	0.3	0.5	0.3	0.4	0.4	0.1	1.0	0.4	0.2	-0.2	0.0	0.1	0.1	0.1	0.5	0.5	0.2	0.1	0.0	0.0	0.1	0.5	0.1	-0.2	0.3	0.0	0.181
0.4	0.6	0.5	0.3	0.5	0.2	0.6	0.6	0.6	0.8	0.7	0.5	0.4	1.0	0.4	0.3	0.1	0.4	0.3	0.2	0.7	0.6	0.6	0.4	0.6	0.2	0.3	0.7	0.3	0.2	0.6	0.5	0.463
0.6	0.4	0.3	0.4	0.6	0.1	0.6	0.4	0.7	0.5	0.4	0.4	0.2	0.4	1.0	0.6	0.3	0.2	0.6	0.0	0.5	0.5	0.3	0.3	0.6	0.3	0.3	0.4	0.4	0.2	0.4	0.6	0.409
0.4	0.5	0.3	0.6	0.3	0.3	0.6	0.3	0.4	0.4	0.3	0.3	-0.2	0.3	0.6	1.0	0.5	0.3	0.6	-0.1	0.4	0.4	0.4	0.3	0.4	0.2	0.0	0.1	0.6	0.2	0.5	0.6	0.342
0.5	0.6	0.1	0.2	0.2	0.4	0.5	0.5	0.6	0.2	0.2	0.1	0.0	0.1	0.3	0.5	1.0	0.3	0.4	0.1	0.3	0.3	0.3	0.3	0.4	0.1	0.1	0.1	0.4	0.2	0.4	0.5	0.298
0.2	0.1	0.4	0.2	0.4	0.3	0.5	0.5	0.1	0.2	0.2	0.6	0.1	0.4	0.2	0.3	0.3	1.0	0.4	0.4	0.4	0.3	0.5	0.3	0.5	0.3	0.1	0.0	0.6	0.4	0.4	0.4	0.313
0.2	0.3	0.3	0.5	0.3	-0.1	0.5	0.4	0.4	0.5	0.2	0.2	0.1	0.3	0.6	0.6	0.4	0.4	1.0	0.1	0.5	0.4	0.5	0.4	0.4	0.2	0.2	0.3	0.5	0.3	0.4	0.3	0.350
0.4	-0.1	0.2	0.3	0.4	0.3	-0.1	0.4	0.0	0.2	0.2	0.3	0.1	0.2	0.0	-0.1	0.1	0.4	0.1	1.0	0.3	0.1	0.3	0.3	0.2	0.0	0.4	0.2	0.3	0.3	0.4	0.1	0.200
0.1	0.7	0.5	0.3	0.5	0.2	0.5	0.7	0.6	0.7	0.7	0.4	0.5	0.7	0.5	0.4	0.3	0.4	0.5	0.3	1.0	0.7	0.7	0.5	0.5	0.1	0.3	0.6	0.5	0.1	0.6	0.3	0.477
0.7	0.4	0.4	0.3	0.5	0.1	0.6	0.6	0.5	0.7	0.8	0.2	0.5	0.6	0.5	0.4	0.3	0.3	0.4	0.1	0.7	1.0	0.6	0.4	0.4	0.1	0.1	0.6	0.4	0.2	0.5	0.3	0.426
0.7	0.5	0.5	0.4	0.4	0.2	0.4	0.6	0.4	0.6	0.5	0.4	0.2	0.6	0.3	0.4	0.3	0.5	0.5	0.3	0.7	0.6	1.0	0.4	0.4	0.4	0.1	0.3	0.5	0.1	0.6	0.4	0.420
0.6	0.4	0.5	0.4	0.4	0.2	0.4	0.7	0.4	0.3	0.3	0.2	0.1	0.4	0.3	0.3	0.3	0.3	0.4	0.3	0.5	0.4	0.4	1.0	0.5	0.2	0.2	0.3	0.4	0.6	0.4	0.5	0.377
0.5	0.5	0.5	0.3	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.0	0.6	0.6	0.4	0.4	0.5	0.4	0.2	0.5	0.4	0.4	0.5	1.0	0.2	0.4	0.3	0.5	0.5	0.4	0.5	0.437
0.4	0.1	0.3	0.0	-0.1	0.1	0.0	0.2	0.2	0.0	0.2	0.5	0.0	0.2	0.3	0.2	0.1	0.3	0.2	0.0	0.1	0.1	0.4	0.2	0.2	1.0	0.4	-0.1	0.2	0.3	0.3	0.1	0.165
0.2	0.3	0.5	0.0	0.1	0.5	0.0	0.2	0.3	0.2	0.4	0.5	0.1	0.3	0.3	0.0	0.1	0.1	0.2	0.4	0.3	0.1	0.1	0.2	0.4	0.4	1.0	0.2	0.1	0.1	0.3	0.0	0.208
0.1	0.3	0.2	0.2	0.6	0.0	0.4	0.6	0.6	0.8	0.7	0.2	0.5	0.7	0.4	0.1	0.1	0.0	0.3	0.2	0.6	0.6	0.3	0.3	0.3	-0.1	0.2	1.0	0.1	0.0	0.4	0.2	0.336
0.5	0.4	0.6	0.5	0.4	0.5	0.5	0.5	0.4	0.4	0.3	0.3	0.1	0.3	0.4	0.6	0.4	0.6	0.5	0.3	0.5	0.4	0.5	0.4	0.5	0.2	0.1	0.1	1.0	0.3	0.6	0.5	0.412
0.4	0.0	0.3	0.2	0.3	0.2	0.2	0.3	0.1	0.1	0.1	0.2	-0.2	0.2	0.2	0.2	0.2	0.4	0.3	0.3	0.1	0.2	0.1	0.6	0.5	0.3	0.1	0.0	0.3	1.0	0.2	0.4	0.228
0.2	0.5	0.3	0.4	0.5	0.2	0.3	0.5	0.4	0.5	0.6	0.3	0.3	0.6	0.4	0.5	0.4	0.4	0.4	0.4	0.6	0.5	0.6	0.4	0.4	0.3	0.3	0.4	0.6	0.2	1.0	0.4	0.408
0.4	0.4	0.3	0.5	0.6	0.3	0.7	0.5	0.4	0.4	0.2	0.3	0.0	0.5	0.6	0.6	0.5	0.4	0.3	0.1	0.3	0.3	0.4	0.5	0.5	0.1	0.0	0.2	0.5	0.4	0.4	1.0	0.376
																																0.363

Table E.3. Inter-participant correlations (Pearson) for ratings of "cadential effectiveness". The column to the right shows the per participant mean correlation, the value at the bottom right shows the overall mean correlation.

Inter-participant Correlations																																			Mean
1.0	0.4	0.5	0.7	0.6	0.7	0.4	0.6	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.4	0.6	0.6	0.5	0.7	0.5	0.6	0.6	0.6	0.4	0.6	0.5	0.5	0.4	0.7	0.5	0.4	0.4	0.563
0.4	1.0	0.5	0.5	0.5	0.5	0.3	0.5	0.4	0.3	0.3	0.4	0.5	0.3	0.4	0.5	0.4	0.1	0.4	0.4	0.4	0.5	0.4	0.5	0.5	0.4	0.3	0.5	0.2	0.4	0.2	0.5	0.3	0.4	0.3	0.399
0.5	0.5	1.0	0.5	0.5	0.5	0.1	0.5	0.4	0.3	0.4	0.3	0.6	0.5	0.4	0.6	0.4	0.4	0.5	0.5	0.3	0.4	0.3	0.4	0.3	0.4	0.3	0.4	0.3	0.3	0.2	0.4	0.3	0.2	0.2	0.391
0.7	0.5	0.5	1.0	0.7	0.7	0.5	0.5	0.7	0.5	0.7	0.5	0.7	0.6	0.7	0.7	0.6	0.3	0.7	0.7	0.5	0.7	0.6	0.6	0.6	0.6	0.4	0.7	0.6	0.5	0.2	0.7	0.6	0.3	0.5	0.568
0.6	0.5	0.5	0.7	1.0	0.7	0.5	0.6	0.6	0.4	0.6	0.6	0.6	0.6	0.7	0.6	0.7	0.2	0.7	0.7	0.5	0.7	0.5	0.7	0.7	0.7	0.4	0.6	0.6	0.6	0.2	0.8	0.5	0.6	0.5	0.573
0.7	0.5	0.5	0.7	0.7	1.0	0.4	0.6	0.7	0.5	0.7	0.5	0.7	0.7	0.7	0.6	0.7	0.3	0.7	0.7	0.6	0.8	0.6	0.7	0.6	0.6	0.3	0.7	0.5	0.6	0.3	0.7	0.7	0.6	0.4	0.600
0.4	0.3	0.1	0.5	0.5	0.4	1.0	0.4	0.3	0.3	0.5	0.3	0.2	0.2	0.3	0.3	0.3	0.2	0.4	0.3	0.2	0.3	0.2	0.3	0.4	0.5	0.3	0.3	0.3	0.2	0.2	0.4	0.2	0.4	0.5	0.328
0.6	0.5	0.5	0.5	0.6	0.6	0.4	1.0	0.5	0.3	0.4	0.6	0.6	0.4	0.5	0.5	0.6	0.2	0.5	0.5	0.5	0.6	0.3	0.5	0.6	0.5	0.3	0.4	0.4	0.5	0.4	0.6	0.3	0.4	0.3	0.472
0.7	0.4	0.4	0.7	0.6	0.7	0.3	0.5	1.0	0.4	0.6	0.4	0.8	0.7	0.7	0.6	0.7	0.2	0.6	0.8	0.6	0.8	0.6	0.7	0.7	0.5	0.3	0.6	0.5	0.4	0.2	0.7	0.8	0.4	0.3	0.551
0.7	0.3	0.3	0.5	0.4	0.5	0.3	0.3	0.4	1.0	0.7	0.4	0.3	0.5	0.5	0.5	0.5	0.4	0.5	0.4	0.5	0.5	0.4	0.5	0.4	0.5	0.3	0.5	0.4	0.4	0.3	0.5	0.5	0.3	0.4	0.446
0.7	0.3	0.4	0.7	0.6	0.7	0.5	0.4	0.6	0.7	1.0	0.5	0.6	0.6	0.5	0.6	0.6	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.6	0.6	0.4	0.6	0.6	0.4	0.3	0.6	0.6	0.5	0.4	0.541
0.6	0.4	0.3	0.5	0.6	0.5	0.3	0.6	0.4	0.4	0.5	1.0	0.4	0.5	0.5	0.6	0.5	0.3	0.5	0.4	0.5	0.6	0.4	0.4	0.6	0.6	0.4	0.5	0.4	0.4	0.4	0.6	0.3	0.4	0.3	0.455
0.6	0.5	0.6	0.7	0.6	0.7	0.2	0.6	0.8	0.3	0.6	0.4	1.0	0.6	0.7	0.6	0.6	0.2	0.6	0.7	0.5	0.8	0.5	0.7	0.6	0.5	0.2	0.6	0.5	0.4	0.1	0.7	0.7	0.4	0.3	0.525
0.6	0.3	0.5	0.6	0.6	0.7	0.2	0.4	0.7	0.5	0.6	0.5	0.6	1.0	0.6	0.7	0.7	0.4	0.7	0.7	0.6	0.7	0.6	0.7	0.7	0.6	0.2	0.6	0.4	0.5	0.2	0.7	0.7	0.5	0.5	0.557
0.6	0.4	0.4	0.7	0.7	0.7	0.3	0.5	0.7	0.5	0.5	0.5	0.7	0.6	1.0	0.6	0.8	0.2	0.7	0.8	0.5	0.8	0.7	0.6	0.6	0.7	0.1	0.7	0.5	0.5	0.2	0.8	0.8	0.5	0.5	0.573
0.6	0.5	0.6	0.7	0.6	0.6	0.3	0.5	0.6	0.5	0.6	0.6	0.6	0.7	0.6	1.0	0.6	0.3	0.7	0.7	0.5	0.6	0.6	0.6	0.6	0.6	0.4	0.6	0.5	0.5	0.3	0.6	0.5	0.4	0.4	0.556
0.6	0.4	0.4	0.6	0.7	0.7	0.3	0.6	0.7	0.5	0.6	0.5	0.6	0.7	0.8	0.6	1.0	0.2	0.6	0.7	0.6	0.8	0.7	0.7	0.7	0.7	0.2	0.8	0.5	0.5	0.2	0.8	0.7	0.5	0.5	0.575
0.4	0.1	0.4	0.3	0.2	0.3	0.2	0.2	0.2	0.4	0.5	0.3	0.2	0.4	0.2	0.3	0.2	1.0	0.3	0.3	0.2	0.3	0.3	0.2	0.3	0.3	0.3	0.3	0.2	0.1	0.2	0.3	0.2	0.1	0.2	0.263
0.6	0.4	0.5	0.7	0.7	0.7	0.4	0.5	0.6	0.5	0.5	0.5	0.6	0.7	0.7	0.7	0.6	0.3	1.0	0.8	0.5	0.7	0.7	0.6	0.6	0.7	0.3	0.6	0.5	0.5	0.1	0.7	0.6	0.4	0.4	0.553
0.6	0.4	0.5	0.7	0.7	0.7	0.3	0.5	0.8	0.4	0.6	0.4	0.7	0.7	0.8	0.7	0.7	0.3	0.8	1.0	0.6	0.8	0.7	0.7	0.7	0.7	0.3	0.7	0.5	0.5	0.2	0.7	0.8	0.4	0.4	0.591
0.5	0.4	0.3	0.5	0.5	0.6	0.2	0.5	0.6	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.6	0.2	0.5	0.6	1.0	0.6	0.5	0.6	0.6	0.5	0.3	0.6	0.4	0.5	0.2	0.6	0.5	0.4	0.2	0.473
0.7	0.5	0.4	0.7	0.7	0.8	0.3	0.6	0.8	0.5	0.6	0.6	0.8	0.7	0.8	0.6	0.8	0.3	0.7	0.8	0.6	1.0	0.8	0.7	0.7	0.7	0.3	0.7	0.6	0.6	0.3	0.7	0.8	0.5	0.5	0.621
0.5	0.4	0.3	0.6	0.5	0.6	0.2	0.3	0.6	0.4	0.5	0.4	0.5	0.6	0.7	0.6	0.7	0.3	0.7	0.7	0.5	0.8	1.0	0.6	0.6	0.6	0.3	0.6	0.5	0.4	0.1	0.6	0.7	0.3	0.4	0.504
0.6	0.5	0.4	0.6	0.7	0.7	0.3	0.5	0.7	0.5	0.6	0.4	0.7	0.7	0.6	0.6	0.7	0.2	0.6	0.7	0.6	0.7	0.6	1.0	0.8	0.6	0.2	0.6	0.5	0.6	0.1	0.7	0.7	0.5	0.3	0.557
0.6	0.5	0.3	0.6	0.7	0.6	0.4	0.6	0.7	0.4	0.6	0.6	0.6	0.7	0.6	0.6	0.7	0.3	0.6	0.7	0.6	0.7	0.6	0.8	1.0	0.7	0.3	0.7	0.4	0.5	0.3	0.8	0.6	0.5	0.5	0.570
0.6	0.4	0.4	0.6	0.7	0.6	0.5	0.5	0.5	0.5	0.6	0.6	0.5	0.6	0.7	0.6	0.7	0.3	0.7	0.7	0.5	0.7	0.6	0.6	0.7	1.0	0.3	0.6	0.5	0.5	0.3	0.7	0.5	0.4	0.5	0.553
0.4	0.3	0.3	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.2	0.2	0.1	0.4	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.3	0.3	1.0	0.2	0.4	0.3	0.3	0.3	0.0	0.1	0.2	0.287
0.6	0.5	0.4	0.7	0.6	0.7	0.3	0.4	0.6	0.5	0.6	0.5	0.6	0.6	0.7	0.6	0.8	0.3	0.6	0.7	0.6	0.7	0.6	0.6	0.7	0.6	0.2	1.0	0.4	0.5	0.2	0.6	0.7	0.4	0.5	0.550
0.5	0.2	0.3	0.6	0.6	0.5	0.3	0.4	0.5	0.4	0.6	0.4	0.5	0.4	0.5	0.5	0.5	0.2	0.5	0.5	0.4	0.6	0.5	0.5	0.4	0.5	0.4	0.4	1.0	0.4	0.2	0.5	0.4	0.4	0.4	0.437
0.5	0.4	0.3	0.5	0.6	0.6	0.2	0.5	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.1	0.5	0.5	0.5	0.6	0.4	0.6	0.5	0.5	0.3	0.5	0.4	1.0	0.4	0.5	0.4	0.4	0.4	0.446
0.4	0.2	0.2	0.2	0.2	0.3	0.2	0.4	0.2	0.3	0.3	0.4	0.1	0.2	0.2	0.3	0.2	0.2	0.1	0.2	0.2	0.3	0.1	0.1	0.3	0.3	0.3	0.2	0.2	0.4	1.0	0.3	0.0	0.2	0.3	0.234
0.7	0.5	0.4	0.7	0.8	0.7	0.4	0.6	0.7	0.5	0.6	0.6	0.7	0.7	0.8	0.6	0.8	0.3	0.7	0.7	0.6	0.7	0.6	0.7	0.8	0.7	0.3	0.6	0.5	0.5	0.3	1.0	0.7	0.5	0.5	0.604
0.5	0.3	0.3	0.6	0.5	0.7	0.2	0.3	0.8	0.5	0.6	0.3	0.7	0.7	0.8	0.5	0.7	0.2	0.6	0.8	0.5	0.8	0.7	0.7	0.6	0.5	0.0	0.7	0.4	0.4	0.0	0.7	1.0	0.5	0.4	0.509
0.4	0.4	0.2	0.3	0.6	0.6	0.4	0.4	0.4	0.3	0.5	0.4	0.4	0.5	0.5	0.4	0.5	0.1	0.4	0.4	0.4	0.5	0.3	0.5	0.5	0.4	0.1	0.4	0.4	0.4	0.2	0.5	0.5	1.0	0.5	0.400
0.4	0.3	0.2	0.5	0.5	0.4	0.5	0.3	0.3	0.4	0.4	0.3	0.3	0.5	0.5	0.4	0.5	0.2	0.4	0.4	0.2	0.5	0.4	0.3	0.5	0.5	0.2	0.5	0.4	0.4	0.3	0.5	0.4	0.5	1.0	0.398
																																			0.492